10.48047/jocaaa.2023.31.03.22

Heterogeneous Graph Transformer based Edge Anomalous Transaction Detection with XAI for Anti-Money Laundering System

Mudit Agarwal

Sr. Vice President, Head of AML and Fraud Compliance Engineering, BNY, New Jersey-08837.
Email: mudit24@yahoo.com

ABSTRACT
Money laundering is a financial crime and it is a critical challenge confronted by the banking sector. Standard Anti-Money Laundering (AML) systems are based on rules or node-based classification, which cannot deal with multi-entity relationship. This work overcomes this issue by presenting an anomalous transaction detection system in edge-level on heterogeneous graphs through Heterogeneous Graph Transformer (HGT) model. Several diversified entities are modelled as heterogeneous graphs and the anomalous transactions can be detected. Finally, Explainable Artificial Intelligence (XAI) technique is included to visualize the attention heatmap for rendering model insight. The performance of the proposed work is justified by comparing it with different baseline models in terms of AUC-ROC, precision-recall and F-score.

Keywords: Anti-Money Laundering, Deep Learning, HGT, anomalous transaction, XAI.

1. Introduction

Money laundering is a persistent hazard to global financial systems, as it promotes illegal activities and destabilizes economic stability and financial institutions. Criminals conceal their illicit financial gains by transforming their stolen funds through money laundering operations, thereby enabling criminal profits from drug trafficking, corruption, and terrorist activities to enter legitimate economic ecosystems [1]. This global challenge poses significant integrity risks to the international financial system, as capital transfers are disrupted and foreign investment decreases. Governments, financial institutions, and regulatory bodies worldwide have prioritized the global fight against money laundering.

Governments established Anti-Money Laundering (AML) regulations in response to the vulnerability of financial systems and the threat of money laundering. An extensive structure of enforcement systems is employed by AML to combat unlawful fund transfers, which combines regulatory frameworks, operational techniques, and technological solutions [2]. AML frameworks are essential for the preservation of international peace by interrupting the availability of terrorist assets and illicit funding, while also providing essential security [3]. The global fight against AML is significantly influenced by the financial sector, particularly banks [4]. Financial institutions are susceptible to illicit activities due to their extensive array of services, which encompasses transaction services, foreign correspondent accounts, and investment platforms.

Despite the fact that money launderers leverage technological advancements, global connections, and banking services to facilitate seamless financial transfers, their detection capabilities are still restricted. The necessity of robust, adjustable AML systems to safeguard financial institutions is illustrated by significant compliance violations. Banks have prioritized the detection of financial crimes. The confirmation of fraud events is contingent upon the use of machine-learning methods in contemporary financial fraud detection [5]. Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL) [6] have recently emerged as game-changing instruments in the fight against AML [7]. In order to identify anomalous activity signals indicative of potential criminal activity, these technology tools extract patterns concealed within immense quantities of transactional data.

Financial institutions can more effectively detect money laundering activities and prevent these unlawful transactions by monitoring transactions in real time, which is facilitated by ML-based algorithms. The algorithms are designed to minimize erroneous incident detection and safeguard against updated illicit activity through dynamic modification. This study is indispensable due to the substantial security dangers that international financial systems face as a result of the ongoing advancement of sophisticated money laundering techniques. The current AML systems are unable to effectively detect or prevent such schemes due to the evolving criminal methods used to conceal illicit transactions. In order to improve efficiency, reduce false positive outcomes, and improve detection accuracy, robust data-driven methods must be

10.48047/jocaaa.2023.31.03.22

developed.

Most of the existing rule-based engines are straightforward and the standard Graph Neural Networks (GNN) consider homogeneous data, where a node is linked with an edge. This could ignore the variations available in the transactions. Hence, this article proposes a system to detect anomalous transaction by modelling the data in a heterogeneous graph with AI explainability. In a heterogeneous graph, multiple entity types with varied relationships are modelled, which makes the problem more realistic. Instead of flagging an account, this work tends to detect every single suspicious transaction. The major work contributions are listed below.

- Edge based anomaly detection is proposed on a heterogeneous graph, which is more realistic.
- Heterogeneous Graph Transformer (HGT) is employed to detect anomalous transactional data.
- AI explainability is included for better data interpretation and understandability.
- Comparing the proposed work with homogeneous GNNs to highlight the qualitative visualizations.

The remainder of this article is organized in the following fashion. Section 2 reviews the related literature concerning anomaly detection in financial transactions, the proposed HGT based anomalous transaction detection model is elaborated in section 3. The performance of the proposed work is evaluated in section 4 and the article is concluded in section 5.

2. Review of Literature

This section intends to review the recent related literature concerning heterogeneous and edge-based data handling.

The work presented in [8] employs HGT to detect fraudulent activities in Bitcoin transactions. However, this work focusses only on crypto networks with minimal real-world labels. In [9], a scalable GNN is presented to detect fraudulent transaction in eBay and this work employed built-in explainability. However, generalization cannot be achieved by this model and meta-path construction is necessary.

In [10], a work to detect money launderers by employing heterogeneous GNNs is presented for AML. Yet, this work involves node-based detection and edge level processing is not its scope. The work presented in [11] proposes a self-explainable heterogeneous transformer for detecting fraudulent activities by using learnable masks. The demerit of this work is its nodal classification and the transaction processing is unclear.

The authors of [12] presented a system to detect multi-step attack through the introduction of edge-based anomaly detection. The major limitation of this work is that it does not focus on financial transactions. In [13], a spectral heterogeneous GNN based on wavelet is presented for anomaly detection. This work involves high computational complexity.

In [14], a meta-path explainability based heterogeneous GNN is presented. Though this work shows better fidelity, it involves node level processing. A graph-based anomaly detection with adaptive message passing is presented in [15]. However, this work focuses on homogeneous graphs with no explainability.

The work proposed in [16] involves heterogeneous GNN for financial-social systems and this work also focusses on nodal processing with lacking explainability. In [17], a dynamic spectral graph anomaly scheme is presented on the basis of spectral GNN, which can handle temporal graphs. However, this work cannot handle heterogeneous data and unexplainable.

The work proposed in [18] presented Anomal-E, which is based on self-supervised edge level GNN for anomaly detection. However, this work falls under the scope of network security. A heterogeneous GNN based fraud detection system is proposed for internet platforms in [19]. However, this work focusses on internet based fraudulent detection and unspecific. Different graph-based anomaly detection schemes are surveyed in [20].

The authors of [21] presents an explainable heterogeneous GNN for land usage graphs and is not under the scope of AML. In [22], a heterogeneous path extraction with link prediction is presented but, it does not focus on anamoly detection.

Inspired by these existing related works, this article attempts to combine three cornerstones that includes edge-level detection, heterogeneous data modelling and explainability under the context of AML.

3. Proposed AML with Edge Anomaly Detection based on HGT

The major goal of this work is to detect anomalous transactions on heterogenous data by employing HGT,

10.48047/jocaaa.2023.31.03.22

in order to enforce strict AML. This work involves four major key phases, which are data acquisition and pre-processing, graph generation, feature engineering, HGT modelling and XAI provisioning. All these steps are described in the following sub-sections.

3.1 Data acquisition and pre-processing

The data acquisition is done from a publicly available PaySim dataset [23], which contains about 63,53,307 records with 11 columns. The columns involved in this dataset are shown in table 1. This section intends to review the recent related literature concerning heterogeneous and edge-based data handling.

Table 1. Columns in the dataset

Column	Description
Step	Time step
Type	Transactional type Cash_out, payment, transfer
Amount	Transactional value
nameorig	Account originator ID of the transaction initiator
Oldbalanceorig	Account balance of originating account before transaction
Newbalanceorig	Account balance of originating account after transaction
namedest	Account destination ID of the transaction receiver
Oldbalancedest	Account balance of destination account before transaction
Newbalancedest	Account balance of destination account after transaction
isfraud	Binary label (1=fraud;0=normal)
Isflaggedfraud	Is the transaction flagged? always 0

This includes mobile transactional data for about a month.

- Data Cleansing: Initially, the data is cleaned to remove any duplicate records and the irrelevant columns are ignored. In this case, the last column "isflaggedfraud" is removed, as it is irrelevant here. The key identifiers of this work are nameorig and namedest, but are changed to numerals to ensure data anonymization.
- Categorical feature encoding: It is an important step here, as this work involves edge-based classification. Hence, the 'type' column is label-encoded, so as to obtain a multi-dimensional sparse vector that indicates the transaction type. The graph is embedded with this categorical representation for processing edge feature set of transactions.
- Feature normalization: The fields with numerals such as 'amount', 'oldbalanceorig', 'newbalanceorig', 'oldbalancedest', 'newbalancedest' are normalized by Z-score normalization for consistent scaling and to stabilize the training process.
- Creation of graph schema: As this work processes heterogeneous data, the dataset is structured by a heterogeneous graph that involves nodes such as account and edges such as 'transfers to'. This edge is associated with a label 'isfraud', in order to carry out supervised learning.
- Class balance management: This dataset shows label imbalance and it is managed by SMOTE by synthetic oversampling. This is to make the dataset with equal fraudulent and normal activities, while avoiding biased detection.

3.2 Graph Generation

The transaction data is converted to a heterogeneous network, where every node and edge represent an entity and relationship/event respectively. The node can be an account, user or device and the edge can be a transaction. To generate a graph, four key steps such as node type identification, edge type definition, attribute assignment to node, edge and graph generation.

- Node type identification: This work considers three nodes such as account, user and device. Here, account denotes sender and receiver accounts that are derived from the source. The nodes user and device are synthesized for the sake of allotting multiple accounts for a user. The device involve a device ID. Hash mapping is then carried out to convert raw identifiers to a numerical node index.
- Edge type definition: The edges involved in this work are 'transfers to', 'owns', 'uses_device'. The financial transaction between two accounts is represented by 'transfers to', which is obtained from the dataset. 'Owns' and 'uses_device' are synthesized, where the first edge indicates the user who owns the account. The 'uses_device' indicates the device used for transaction.
- Attribute assignment to nodes and edges: The attributes assigned for nodes and edges are shown in table 2.

10.48047/jocaaa.2023.31.03.22

Table 2. Attributes of Nodes and Edges

Entity	Attributes
Account (Node)	Transactions count, ratio of fraudulent activity, average balance
User (Node)	Total accounts, behavioural status
Device (Node)	Usage frequency, rate of reuse
Transfers_to (Edge)	Scaled amount, type, average balance

- Graph generation: The graph is then generated which can capture the diversified entities and relations through heterogeneity and can detect fraudulent transactions effectively. Let the nodes and edges be represented by

$$Nodes = \begin{cases} Account: \{A_1, A_2, \dots, A_n\} \\ User: \{U_1, U_2, \dots, U_n\} \\ Device: \{D_1, D_2, \dots, D_n\} \end{cases} \quad (1)$$

$$Edges = \begin{cases} Transfers_{to}: A_1 \rightarrow A_2 \\ Owns: U_1 \rightarrow A_1 \\ User\ Device: \{A_1 \rightarrow D_1\} \end{cases} \quad (2)$$

Hence, a graph with nodes and edges are generated with their associations.

3.3 Feature Engineering

The objective of feature engineering is to convert a raw transactional data to a meaningful nodal and edge-based features, such that behavioural patterns of users, semantics of the transaction and the relationship between networks are explored. These features form the base for HGT's learning ability to detect anomalies. The feature engineering is carried out on nodes, edges and then normalized. The features of nodes are presented as follows.

$$Nodes = \begin{cases} Account\ \{C_t, Transamt_{avg}, Bal_{avg}, CF_{net}\} \\ User\{Acc_{Tot}, Transvol_{tot}, D_{count}\} \\ Node: Device\{Ruse_{count}\} \end{cases} \quad (3)$$

Where $C_t, Transamt_{avg}, Bal_{avg}, CF_{net}$ denote count of transactions, average transaction amount, average balance and net flow of the account. $Acc_{Tot}, Transvol_{tot}, D_{count}$ represent the total number of accounts associated with the user, summation of all transactions by all the owned accounts and count of utilized unique devices respectively. $Ruse_{count}$ indicate the count of accounts that utilize the same device. The edge features of type 'transfers_to' involves basic and computed edge features as in eqns. (4) and (5).

$$Edges = Transfers_{to}\ R\{Amt, type, step\} \quad (4)$$

$$Edges = Transfers_{to} C\{bal_{orig}, bal_{dest}, amt_{scld}, amt_{ratio}\} \quad (5)$$

In the above equations, $Amt, type, step$ indicates the transaction value, transaction type and hourly step of transaction. The bal_{orig} and bal_{dest} are computed by

$$bal_{orig} = Oldbalanceorig - Newbalanceorig \quad (6)$$

$$bal_{dest} = Oldbalancedest - Newbalancedest \quad (7)$$

amt_{scld}, amt_{ratio} are the normalized amount using z-score normalization technique and ratio of the sender's previous and the current balance.

Hence, the feature engineering helps to enhance the anomaly detection and supports HGT to learn from different nodes and edges.

3.4 HGT Modelling

This HDT DL model learns from heterogeneous graphs with several nodes and edges. Edge level classification is performed to detect anomalous transactions and explainable outcomes are returned by attention weights. During the training stage, the HGT is passed with inputs such as nodes, edges, nodal features, edge features and edge labels. Every layer of HGT performs three activities such as type-aware attention, message aggregation and node embedding update. The parameters used for learning are tabulated in table 3.

Table 3. Training parameters

10.48047/jocaaa.2023.31.03.22

Parameter	Value
Optimizer	Adam
Learning rate	1e-3 to 1e-4
Loss function	Binary cross entropy
Sampling strategy	Minibatches (edge-wise)

The attention weights are computed for every (node_type, edge_type) pair.

$$Attention_{(i,j,r)} = softmax(Q_i^T \cdot R^r . K_j) \quad (8)$$

Where target node's query, source node's key and relation-based matrix for edge type r are denoted by Q_i^T, K_j, R^r respectively. The type_specific feed forward networks are updated. For every edge, the source and destination node embeddings are combined, which is represented by $h_{edge} = [h_{src} \oplus h_{dest}]$. The anomaly is detected by passing it to Multilayer Perceptron (MLP).

$$y_{edge} = \sigma\left(W_2 \cdot ReLU\left(W_1 . h_{edge}\right)\right) \quad (9)$$

Hence, the edge-based anomalous transactions are classified and the explainability of the outcome is presented in the next section.

3.5 XAI provisioning

HGT is explainable on its own with the help of built-in attention scores through which the type of neighbour who is responsible for the decision taken and the relationship between them can be visualized. The heatmaps over neighbourhood is generated and a sample image is presented in figure 1.

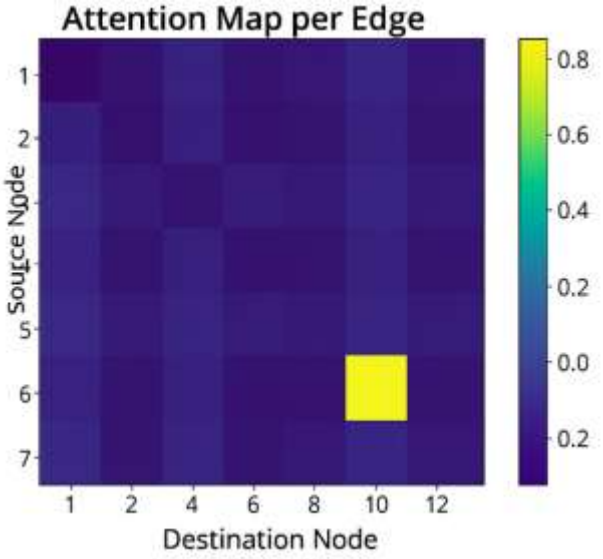


Figure 1. Sample attention map per edge

In the above figure, x and y axes represent the destination and source nodes. Each cell indicates the attention rendered by the destination node to the source node. In this case, node 10 shows more attention towards node 6, which might fall under three varied scenarios. Node 6 may involve in recent money transaction to node 10 or node 6 is linked through a device or user or node 6 may be fraudulent. In all these cases, nodes 6 and 10 are closely related.

4. Results and Discussion

The proposed work is implemented on a stand-alone computer with 24 GB RAM with Python. The train/test ratio of the proposed work is 80/20, which is 2300 and 1000 records respectively. The performance of this work is analysed by employing standard performance metrics such as precision, recall, F-measure, AUC-ROC and precision-recall curve. The performance of the proposed HGT based model is compared against Graph Convolutional Networks (GCN), Hierarchical Attention Network (HAN), Relational GCN (R-GCN) and rule-based AML in terms of precision, recall and F-measure rates, as shown in figure 2. The confusion matrix is shown in Figure 3.

10.48047/jocaaa.2023.31.03.22

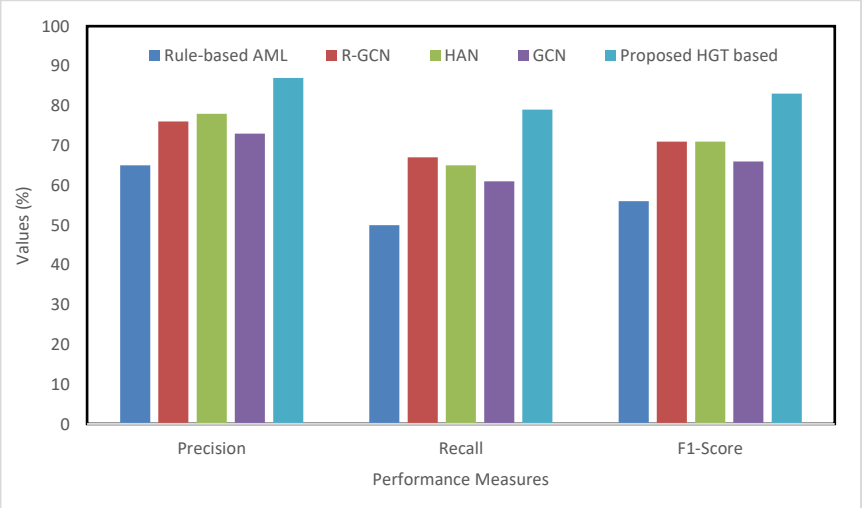


Figure 2. Performance comparison

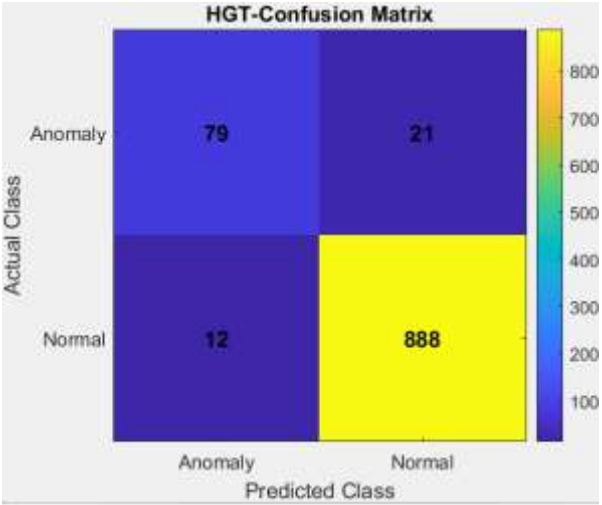
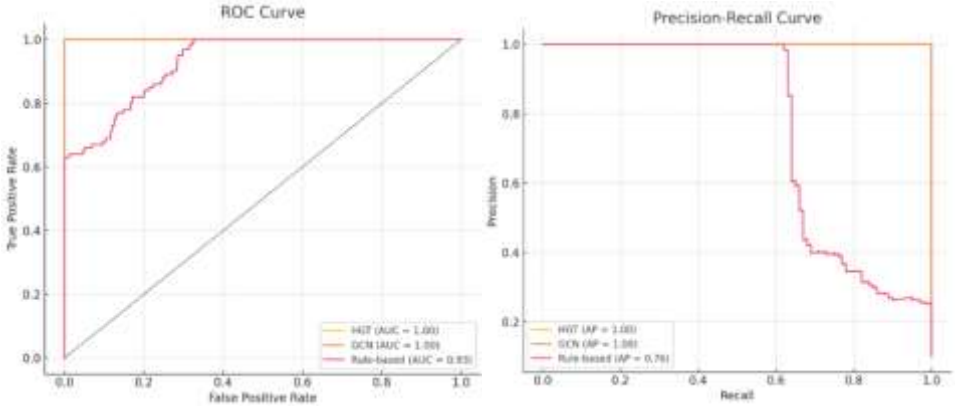


Figure 3. Confusion matrix



(a) (b)

Figure 4. (a) ROC Curve (b) Precision-Recall curve

From the experimental results, it is evident that the HGT performs better when compared to other techniques. The explainable outcomes make it more interpretable and understandable on-the-go. In addition, the false alarms are minimized and is highly suitable for anomaly detection to enforce AML.

5. CONCLUSION

This article presents a HGT based explainable model for anomalous transaction detection to enforce AML. This work considers the real-world transactional data by considering heterogeneous graph that is

10.48047/jocaaa.2023.31.03.22

comprised of numerous entities and it focusses on edge-level anomaly detection, which means the fraudulent transactions are detected. Incorporation of feature engineering and model optimization helps the proposed HGT model to learn multi-relational dependencies. The built-in attention mechanism is utilized to attain explainability and interpretability of the model. The performance of the model is quite satisfactory, when compared to the existing models. In future, temporal behavioural tracking of the user can be included and federated learning can be employed for cross-bank collaboration.

References

1. ICAI, (2015). "Importance of Anti-Money Laundering Measures and Effective KYC in Financial Transactions." kb.icai.org. Available at: https://kb.icai.org/pdfs/PDFFile5b28c97d06d877.47667992.pdf (Accessed: 8 April 2025).
2. Le Khac, N. A., & Kechadi, M. T. (2010). "Application of Data Mining for Anti-Money Laundering Detection: A Case Study." Proceedings -IEEE Int. Conf. Data Mining, ICDM, 577–584. https://doi.org/10.1109/ICDMW.2010.66.
3. Naheem, M. A. (2019). "Anti-Money Laundering/Trade-Based Money Laundering Risk Assessment Strategies – Action or Re-action Focused?" Journal of Money Laundering Control. https://doi.org/10.1108/JMLC-01-2016-0006.
4. Viritha, B., Mariappan, V., & Venkatachalapathy, V. (2015). "Combating Money Laundering by the Banks in India: Compliance and Challenges." Journal of Investment Compliance, 16(4): pp. 78–95. https://doi.org/10.1108/JOIC-07-2015-0044.
5. Sadgali, I., Sael, N., & Benabbou, F. (2019). "Performance of Machine Learning Techniques in the Detection of Financial Frauds." Procedia Computer Science, 148: pp. 45–54. https://doi.org/10.1016/j.procs.2019.01.007.
6. Kolluri, V. (2015). "A Comprehensive Analysis on Explainable and Ethical Machine: Demystifying Advances in Artificial Intelligence." International Research Journal, 2(7).
7. Zantalis, F., Koulouras, G., Karabetsos, S., & Kandris, D. (2019). "A Review of Machine Learning and IoT in Smart Transportat ion." Future Internet, 11(4): pp. 1–23. https://doi.org/10.3390/FI11040094.
8. Liu, Y., Pan, S., Wang, Y. G., Xiong, F., Wang, L., Chen, Q., & Lee, V. C. (2021). Anomaly detection in dynamic graphs via transformer. *IEEE Transactions on Knowledge and Data Engineering*, *35*(12), 12081-12094.
9. Rao, S. X., Zhang, S., Han, Z., Zhang, Z., Min, W., Chen, Z., ... & Zhang, C. (2020). xFraud: explainable fraud transaction detection. arXiv preprint arXiv:2011.12193.
10. Johannessen, F., & Jullum, M. (2023). Finding money launderers using heterogeneous graph neural networks. arXiv preprint arXiv:2307.13499.
11. Qin, Z., Liu, Y., He, Q., & Ao, X. (2022, October). Explainable graph-based fraud detection via neural meta-graph search. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 4414-4418).
12. Mao, B., Liu, J., Lai, Y., & Sun, M. (2021). MIF: A multi-step attack scenario reconstruction and attack chains extraction method based on multi-information fusion. *Computer Networks*, *198*, 108340.
13. Dehghani, M., Niknam, T., Ghiasi, M., Bayati, N., & Savaghebi, M. (2021). Cyber-attack detection in dc microgrids based on deep machine learning and wavelet singular values approach. *Electronics*, *10*(16), 1914.
14. Li, Y., Liu, L., Wang, G., Du, Y., & Chen, P. (2022). EGNN: Constructing explainable graph neural networks via knowledge distillation. *Knowledge-Based Systems*, *241*, 108345.
15. Chen, B., Zhang, J., Zhang, X., Dong, Y., Song, J., Zhang, P., ... & Tang, J. (2022). Gccad: Graph contrastive coding for anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, *35*(8), 8037-8051.
16. Cardoso, M., Saleiro, P., & Bizarro, P. (2022, November). Laundrograph: Self-supervised graph representation learning for anti-money laundering. In *Proceedings of the third ACM international conference on AI in finance* (pp. 130-138).
17. Liu, Y., Pan, S., Wang, Y. G., Xiong, F., Wang, L., Chen, Q., & Lee, V. C. (2021). Anomaly detection in dynamic graphs via transformer. *IEEE Transactions on Knowledge and Data Engineering*, *35*(12), 12081-12094.
18. Caville, E., Lo, W. W., Layeghy, S., & Portmann, M. (2022). Anomal-E: A self-supervised network intrusion detection system based on graph neural networks. Knowledge-based systems, 258, 110030.
19. Tang, S., Jin, L., & Cheng, F. (2021). Fraud detection in online product review systems via heterogeneous graph transformer. Ieee Access, 9, 167364-167373.
20. Kim, H., Lee, B. S., Shin, W. Y., & Lim, S. (2022). Graph anomaly detection with graph neural networks:

10.48047/jocaaa.2023.31.03.22

Current status and challenges. IEEE Access, 10, 111820-111829.
21. Aviyente, S., & Karaaslanli, A. (2022). Explainability in Graph Data Science: Interpretability, replicability, and reproducibility of community detection. *IEEE Signal Processing Magazine*, *39*(4), 25-39.
22. Zhang, S., Zhang, J., Song, X., Adeshina, S., Zheng, D., Faloutsos, C., & Sun, Y. (2023, April). PaGE-Link: Path-based graph neural network explanation for heterogeneous link prediction. In Proceedings of the ACM Web Conference 2023 (pp. 3784-3793).
23. https://www.kaggle.com/datasets/ealaxi/paysim1

Author Biography

Mudit Agarwal is a seasoned technology leader and innovator, currently serving as Head of AML (Anti-Money Laundering) and Fraud Engineering Compliance at BNY Mellon. With over two decades of experience in financial technology, the author has led transformative initiatives that advance risk management, financial crime prevention, and operational resilience. He is widely recognized for his work in modernizing compliance frameworks using cloud computing and real-time data analytics. His leadership was instrumental in deploying the industry's first payment application in a public cloud. A passionate advocate and a winner of an international blockchain hackathon, and continues to inspire innovation through mentorship and thought leadership in fraud detection, AML systems, and secure payment processing. He holds a strong academic and engineering background and currently facilitates the secure processing of over $2.5 trillion in financial messages daily through BNY Mellon's infrastructure.