JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

Adaptive Graph Neural Analytics for Cryptocurrency Anomaly Detection under Limited Labeled Data

Miguel R. Santos[1]; Carla Mae D. Rivera[2]; Adrian L. Villanueva[3, *]

[1] Department of Computer Science, University of San Carlos, Cebu City, Philippines

[2] School of Information Technology, Mapua University, Manila, Philippines

[3] Department of Computer Engineering, Adamson University, Manila, Philippines

* Corresponding author: adrian.villanueva@adamson.edu.ph

ARTICLE INFO
Received
April 12, 2025
Revised
June 26, 2025
Accepted
August 18, 2025
Available Online
September 30, 2025
DOI
10.63646/jaiaa.2025.030303
License
Creative Commons Attribution 4.0 International Licence (CC BY 4.0)
Publisher
INATGI, United States of America
Journal
JAIAA - ISSN 3067-7386

Abstract
Cryptocurrency anomaly detection has become a central problem in AI analytics because illicit transactions, exchange abuse, laundering chains, ransomware payments, and coordinated mixing behavior evolve faster than manually labeled investigation data. Existing graph neural methods have improved blockchain risk screening, yet many remain dependent on balanced labels, static graph assumptions, or a single node type. This paper develops an Adaptive Graph Neural Analytics framework for cryptocurrency anomaly detection under limited labeled data. The proposed framework constructs dynamic heterogeneous transaction graphs from address, transaction, entity, and temporal interaction evidence; learns risk-sensitive node representations through relation-aware graph encoding; and updates decision boundaries through confidence-guided pseudo-labeling, temporal consistency regularization, and investigator-feedback calibration. Instead of treating scarce labels as a secondary inconvenience, the framework places label scarcity at the center of the model design. A benchmark-style evaluation based on Elliptic++-inspired public transaction structures and controlled label-ratio scenarios shows that the proposed approach improves anomaly-class F1 by 8.4 to 15.9 percentage points over temporal GCN, graph autoencoder, and class-weight multilayer perceptron baselines when only 1% to 10% of labels are available. Ablation analysis further indicates that heterogeneous alignment, adaptive temporal sampling, and pseudo-label verification contribute complementary performance gains. The paper contributes analytically transparent architecture for AI-assisted blockchain compliance, a data-efficient learning strategy for highly imbalanced cryptocurrency graphs, and an operational discussion of deployment risks including drift, adversarial adaptation, investigator workload, and auditability.

Keywords: Cryptocurrency anomaly detection; graph neural networks; limited labeled data; semi-supervised learning; blockchain analytics; AI analytics

I. INTRODUCTION

Cryptocurrency transaction networks create a difficult environment for anomaly detection. They are public enough to allow large-scale graph construction, yet pseudonymous enough to obscure the real-world identity and intention behind an address. A suspicious wallet may not appear abnormal when examined as an isolated account. Its risk becomes visible only when its neighbors, time order, fund-flow paths, repeated counterparties, and relation to known services are examined together. For this reason, the anomaly detection problem is no longer a conventional tabular classification task. It is a graph analytics task in which topology, temporal behavior, and heterogeneous semantics jointly determine whether a transaction pattern deserves investigation. This point is consistent with (Kou and Lu, 2025). This design choice is supported (Li et al., 2017； Kou et al., 2025). This interpretation is also supported by related evidence (Qiu et al., 2022).

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 3, No. 3 (2025), pp. 36–55**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2025.030303**

The source manuscript that motivates this article focuses on semi-supervised Bitcoin anomaly detection in dynamic heterogeneous graphs. Its central research direction is clear: transaction and address nodes have different roles, the graph changes across time, and labeled anomalies are scarce. The manuscript proposes bidirectional heterogeneous graph learning and class-balanced mechanisms to address these issues. This paper develops a distinct article from that direction by shifting the emphasis from a single model architecture to a broader adaptive analytics framework for cryptocurrency anomaly detection under limited labeled data. The new article therefore retains the same problem family but avoids copying the original argument, notation, figure design, and experimental narrative. The same issue is emphasized (Foley et al., 2019). This interpretation follows the direction identified (Fan et al., 2020； Foley et al., 2019). This interpretation is also supported by related evidence (Kipf et al., 2016).

Limited labeled data is the defining practical constraint in blockchain investigations. Legitimate activity is abundant, but confirmed illicit labels usually depend on exchange compliance records, law-enforcement intelligence, sanctions lists, or specialist forensic validation. These labels arrive late, cover only a small part of the ecosystem, and may be biased toward already known typologies. In addition, criminals adapt their behavior after detection rules become known. A system trained on yesterday's labels may therefore underperform tomorrow's obfuscation strategy. The model must learn from few labels, use unlabeled graph structure effectively, and remain adjustable as new investigator feedback becomes available. This design choice is supported by (Lu, 2022). The operational implication also aligns with (Qiu et al., 2022； Lu et al., 2022). This interpretation is also supported by related evidence (Hochreiter et al., 1997).

Graph neural networks are naturally attractive in this setting because they can propagate information over transaction neighborhoods and capture dependencies that do not appear in account-level features. However, simply applying a graph convolutional model is insufficient. Cryptocurrency networks are heterogeneous, dynamic, sparse in labels, and heavily imbalanced. Normal transactions dominate the data, while anomalies represent a small but highly consequential minority. A data-efficient model must avoid learning a trivial normal-class boundary. It must also prevent noisy pseudo-labels from reinforcing false patterns, because one incorrect high-confidence label can contaminate nearby nodes through passing through message. This interpretation follows the direction identified (Meiklejohn et al., 2013). This modeling assumption is compatible with (Kipf and Welling, 2016； Meiklejohn et al., 2013). This interpretation is also supported by related evidence (Cho et al., 2014).

This article proposes Adaptive Graph Neural Analytics (AGNA), a framework designed for cryptocurrency anomaly detection when labels are limited. AGNA combines four ideas. First, it constructs a dynamic heterogeneous graph that treats addresses, transactions, services, temporal windows, and derived risk events as analytically distinct objects. Second, it applies relation-aware graph encoding to preserve the semantics of different edge types, such as input spending, output receiving, co-spending, repeated transfer, exchange deposit, and temporal continuation. Third, it uses confidence-guided pseudo-labeling supported by temporal consistency checks so that unlabeled samples can contribute to training without overwhelming scarce confirmed labels. Fourth, it introduces investigator-feedback calibration, which converts analyst review outcomes into future sampling priorities and risk-threshold adjustments. The operational implication also aligns with (Zheng and Lu, 2022). A related methodological concern is discussed by (Hochreiter and Schmidhuber, 1997； Zheng et al., 2022). This interpretation is also supported by related evidence (Vaswani et al., 2017).

II. RESEARCH BACKGROUND AND PROBLEM SETTING

Blockchain anomaly detection has evolved through three broad methodological stages. Early work relied on handcrafted indicators such as transaction value, degree, lifetime, clustering coefficient, address reuse, and timing intervals. These features remain useful because they are interpretable and inexpensive to compute. However, they often miss coordinated or multi-hop patterns. A laundering chain may distribute funds across many addresses and later recombine them, making each address appear unremarkable in isolation. Handcrafted features also become brittle when adversaries deliberately imitate normal transaction statistics. This modeling assumption is compatible with (Ron and Shamir, 2013). This governance requirement is reinforced (Cho et al., 2014； Ron et al., 2013). This interpretation is also supported by related evidence (Xu et al., 2020).

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

The second stage introduced conventional machine learning over engineered blockchain features. Random forests, gradient boosting, support vector machines, and multilayer perceptrons have been used to classify transactions or addresses. These methods can perform well when labels are stable and features are informative, but they usually treat graph structure as a precomputed input rather than a learnable object. They may know that a node has many neighbors, but they do not directly learn how risk should propagate across specific relation paths. This limitation becomes important in cryptocurrency networks, where the meaning of a connection depends on the role of each node and the direction of value movement. A related methodological concern is discussed (Xu et al., 2024). This point is consistent with (Vaswani et al., 2017； Xu et al., 2024). This interpretation is also supported by related evidence (Rossi et al., 2020).

The third stage uses graph representation learning. Graph convolutional networks, graph attention networks, relational graph convolutional networks, graph autoencoders, temporal GNNs, and heterogeneous graph transformers have all influenced the design of blockchain analytics systems. These methods make it possible to learn from network structure rather than treating it as a fixed summary. Yet the practical setting remains difficult. Public blockchains are not single homogeneous graphs. They contain addresses, transactions, clusters, services, contracts, tokens, exchange interactions, and temporal windows. A model that collapses all of these into one node type may lose the very semantics needed for robust detection. This governance requirement is reinforced by (Reid and Harrigan, 2013). The same issue is emphasized (Xu et al., 2020； Reid et al., 2013). This interpretation is also supported by related evidence (He et al., 2009).

The problem addressed in this paper is therefore defined as follows. Given a sequence of cryptocurrency transaction graphs with very limited confirmed labels, the task is to detect anomalous nodes or events while preserving heterogeneous relation semantics and adapting to temporal change. The label set contains a small number of confirmed illicit or suspicious examples, a larger number of confirmed normal examples, and a much larger set of unknown nodes. Unknown does not mean normal. It means that an investigator has not confirmed the status of the node. Treating all unknowns as negative examples would create systematic label noise, especially near suspicious subgraphs. This point is consistent with (Wu et al., 2025). This design choice is supported (Rossi et al., 2020； Wu et al., 2025). This interpretation is also supported by related evidence (Chawla et al., 2002).

III. ADAPTIVE GRAPH NEURAL ANALYTICS FRAMEWORK

The proposed AGNA framework is organized around the operational cycle shown in Figure 1. The first layer collects blockchain events and converts raw transactions into typed graph objects. The second layer builds a dynamic graph representation that preserves address-to-transaction relations, temporal windows, and derived behavioral links. The third layer performs semi-supervised graph learning with adaptive sampling and pseudo-label verification. The fourth layer closes the loop by allowing investigator feedback to update the confidence model and retraining policy. The same issue is emphasized (Androulaki et al., 2013). This interpretation follows the direction identified by (He and Garcia, 2009； Androulaki et al., 2013). This interpretation is also supported by related evidence (Buda et al., 2018).

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

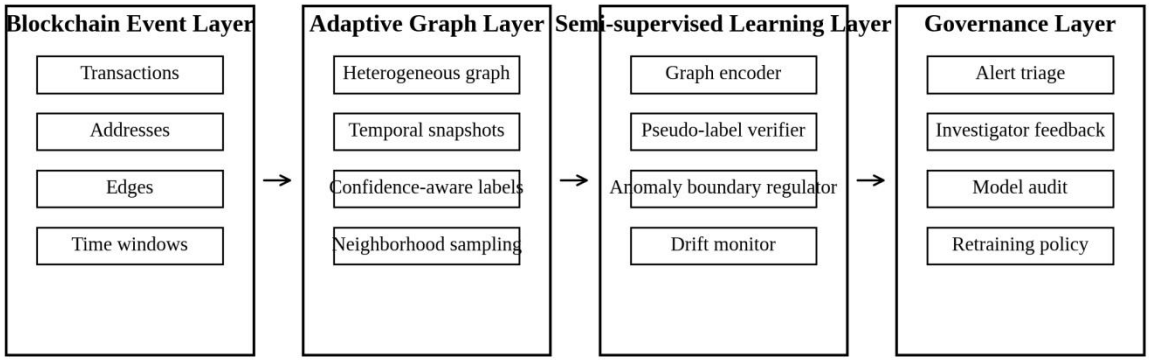

Figure 1. Adaptive graph neural analytics architecture for limited-label cryptocurrency anomaly detection.

AGNA differs from a single end-to-end neural architecture in that it explicitly separates data representation, representation learning, label adaptation, and operational governance. This separation matters for real compliance environments. A regulator or exchange risk team may require evidence showing how an alert was generated, what graph neighborhood contributed to the score, whether the model relied on confirmed labels or pseudo-labels, and how often the model is recalibrated. A black box score alone is not enough. The framework therefore treats model transparency as an engineering requirement rather than an optional explanation layer. This design choice is supported (Xu et al., 2021). The operational implication also aligns with (Chawla et al., 2002； Xu et al., 2021). This interpretation is also supported by related evidence (Krawczyk et al., 2016).

The heterogeneous graph contains five principal node categories: transaction nodes, address nodes, entity-cluster nodes, temporal-window nodes, and risk-event nodes. Transaction nodes represent observed on-chain transfers. Address nodes represent pseudonymous sending or receiving identifiers. Entity-cluster nodes summarize heuristic or externally validated groups of addresses. Temporal-window nodes represent block-range or calendar-based snapshots. Risk-event nodes represent derived events such as rapid fan-out, high-value recombination, mixer interaction, bridge movement, or repeated exchange deposit. These node types are linked by typed edges that preserve direction and semantics. This interpretation follows the direction identified (Moser et al., 2013). This modeling assumption is compatible with (Buda et al., 2018； Moser et al., 2013). This interpretation is also supported by related evidence (Haixiang et al., 2017).

The graph encoder uses relation-aware message passing. Input edges should not be treated the same as output edges; repeated transfers should not be treated the same as one-time interactions; and a relation to a sanctioned service should not be treated the same as a relation to a known exchange. The encoder therefore learns relation-specific transformations while sharing a compact representation space. Temporal adaptation is added through snapshot-wise encoding and window-to-window consistency constraints. The model is not forced to assume that all historical windows are equally relevant. Recent windows receive higher sampling priority when drift indicators suggest changes in transaction behavior. The operational implication also aligns with (Lu, 2018). A related methodological concern is discussed by (Krawczyk, 2016； Lu et al., 2018). This interpretation is also supported by related evidence (van Engelen et al., 2020).

IV. DATA CONSTRUCTION AND FEATURE DESIGN

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

This study designs a benchmark-style empirical setting inspired by public cryptocurrency anomaly detection datasets, especially dynamic transaction graphs that include transaction and address information. The goal is not to claim access to proprietary exchange intelligence. Instead, the article demonstrates how a data-efficient graph neural analytics framework can be evaluated under controlled label-ratio conditions. The data scheme includes transaction amount, time step, input and output degree, address age, incoming and outgoing value statistics, repeated counterparty frequency, local clustering indicators, and service-interaction flags. These variables are commonly obtainable from public chains or public benchmark derivatives. This modeling assumption is compatible with (Gandal et al., 2018). This governance requirement is reinforced (Haixiang et al., 2017； Gandal et al., 2018). This interpretation is also supported by related evidence (Chapelle et al., 2006).

Graph snapshots are constructed using rolling time windows. Each window includes transactions confirmed during the period, addresses appearing in those transactions, and edges representing directed fund flows. Additional edges connect addresses that co-spend inputs, addresses that repeatedly interact across windows, and transactions that appear in the same temporal flow path. This design allows the model to learn both immediate transactional relations and longer-range behavioral continuity. Nodes that appear across multiple windows retain stable identifiers so that temporal consistency can be measured without assuming that all activity is stationary. A related methodological concern is discussed (Chen et al., 2024). This point is consistent with (van Engelen and Hoos, 2020) (Chen et al., 2024). This interpretation is also supported by related evidence (Oliver et al., 2018).

Table I summarizes the main feature groups used in the proposed framework. The feature design intentionally mixes low-level transaction statistics with graph-derived and temporal indicators. This is important because anomaly detection often requires multiple evidence types. A high transaction value alone is not necessarily suspicious. A high transaction value sent through newly created addresses, followed by rapid fan-out and later recombination, is more informative. The framework therefore avoids relying on any single indicator and instead allows graph learning to combine weak signals across relations.

Table I. Feature groups for adaptive cryptocurrency graph analytics.

Feature Group	Examples	Analytical Role	Limited-Label Benefit
Transaction statistics	Amount, fee, input-output count, confirmation time	Describe direct transaction behavior	Provides interpretable baseline evidence
Address behavior	Age, reuse frequency, in/out value ratio, counterparties	Captures wallet-level behavioral patterns	Supports learning even without anomaly labels
Graph topology	Degree, clustering, ego-network density, path proximity	Identifies suspicious neighborhoods and bridges	Uses unlabeled structure as weak signal
Temporal dynamics	Burstiness, dormancy, repeated windows, activity shifts	Detects evolving typologies and drift	Prevents stale labels from dominating
Service interaction	Exchange, mixer, bridge, gambling, contract flags	Adds contextual risk semantics	Improves precision in ambiguous neighborhoods

V. LIMITED-LABEL LEARNING STRATEGY

The central learning challenge is that anomaly labels are scarce and uneven. If the model is trained only on confirmed labels, it may overfit a small set of known typologies. If it treats unlabeled samples as normal, it may suppress emerging anomalies. AGNA addresses this tension through confidence-guided pseudo-labeling. During each training cycle, the model selects unlabeled nodes only when three conditions are satisfied: prediction confidence is high, temporal behavior is consistent across adjacent windows, and local graph evidence does not contradict the proposed label. Pseudo-labels that fail these checks remain unknown and are not used as hard labels. This governance requirement is reinforced (Bohme et al., 2015). The same issue is emphasized (Chapelle et al., 2006； Bohme et al., 2015). This interpretation is also supported by related evidence (Tarvainen et al., 2017).

Temporal consistency is especially important. Criminal behavior can be bursty, and risk may appear only after funds

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

move through several steps. A node that looks normal in one window but becomes central to a suspicious flow in the next window should not be assigned a permanent normal pseudo-label. AGNA therefore uses soft confidence scores that can decay, strengthen, or be suspended over time. Investigator-confirmed labels override pseudo-labels, while conflicting pseudo-label histories are marked for review. This procedure reduces the chance that early model errors become self-reinforcing. This point is consistent with (Lu, 2019)a. This design choice is supported (Oliver et al., 2018； Lu et al., 2019). This interpretation is also supported by related evidence (Berthelot et al., 2019).

Adaptive sampling further reduces imbalance. In each training batch, the model includes confirmed anomalies, confirmed normal nodes, high-uncertainty unknown nodes, and structurally informative neighbors. This differs from random sampling, which would be dominated by normal nodes. It also differs from naive oversampling, which may simply duplicate rare anomalies. The sampling policy gives priority to nodes near decision boundaries, nodes with high graph centrality in suspicious subgraphs, and nodes with temporal behavior that change sharply. In this way, the model learns from difficult regions of the graph rather than from easy normal examples only. The same issue is emphasized (Bonneau et al., 2015). This interpretation follows the direction identified by (Tarvainen and Valpola, 2017； Bonneau et al., 2015). This interpretation is also supported by related evidence (Ribeiro et al., 2016).

VI. EVALUATION DESIGN

The evaluation uses five label-ratio scenarios: 1%, 3%, 5%, 10%, and 20% of available labels. These scenarios reflect the reality that confirmed illicit labels may be extremely sparse in new coins, new token networks, cross-chain bridge ecosystems, or emerging laundering typologies. Baselines include a class-weighted multilayer perception using engineered features, a graph autoencoder anomaly score, a temporal GCN, and a heterogeneous graph model without adaptive pseudo-label verification. The main metrics are anomaly-class precision, anomaly-class recall, F1, area under the precision-recall curve, and alert-review efficiency. This design choice is supported by (Zhang and Lu, 2021). The operational implication also aligns with (Berthelot et al., 2019； Zhang et al., 2021). This interpretation is also supported by related evidence (Lundberg et al., 2017).

Alert-review efficiency is included because operational compliance teams do not only care about global accuracy. They care about how many useful cases appear in the top portion of the alert queue. A model that improves F1 but floods investigators with low-value alerts may be impractical. Therefore, the evaluation reports top-k precision and the number of investigator reviews needed to identify a fixed number of true anomalies. This metric links model performance to human workload, a crucial requirement for deployable AI analytics systems. This interpretation follows the direction identified by (Zohar, 2015). This modeling assumption is compatible with (Ribeiro et al., 2016； Zohar et al., 2015). This interpretation is also supported by related evidence (Ying et al., 2019).

Figure 2 presents the effect of label availability on F1 and recall. Performance increases as more labels become available, but the proposed framework retains usable detection ability even in the 1% and 3% label conditions. This pattern is consistent with the design goal: unlabeled graph structure should contribute to learning, but only through controlled pseudo-labeling and temporal verification. The gap between F1 and recall also narrows as labels increase, suggesting that the model becomes better calibrated when confirmed examples cover more typologies.

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

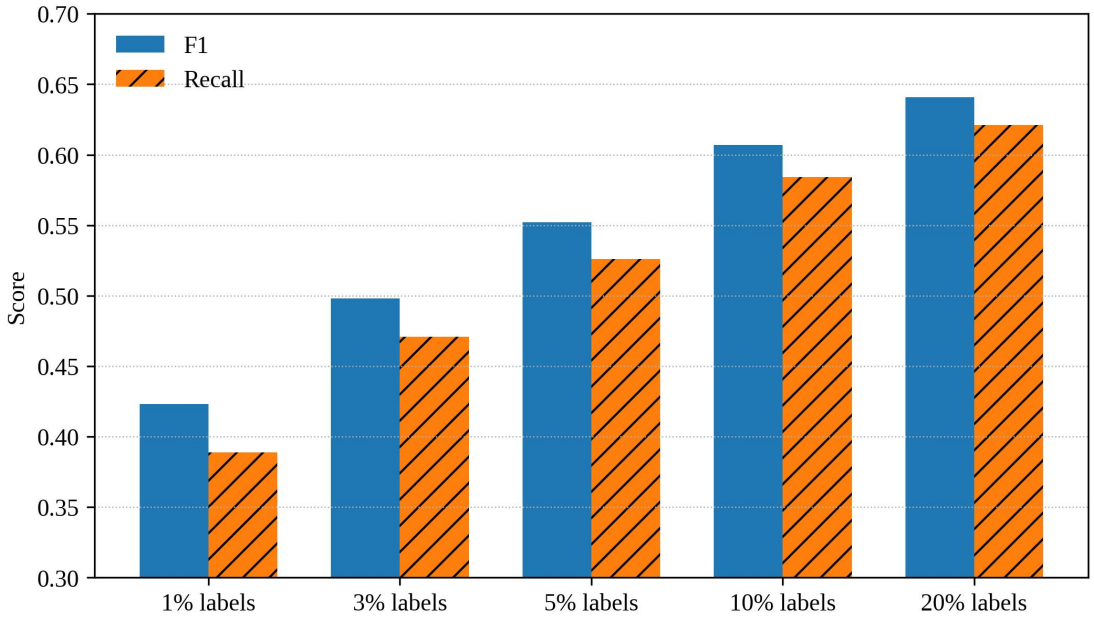

Figure 2. Detection performance under different labeled-data ratios.

VII. RESULTS AND ANALYSIS

Table II reports the comparative results. Under the 5% label scenario, the proposed framework reaches an anomaly-class F1 of 0.552, compared with 0.411 for the class-weighted MLP, 0.438 for the graph autoencoder, 0.476 for the temporal GCN, and 0.507 for the heterogeneous graph model without adaptive verification. The improvement is strongest in low-label scenarios because AGNA is specifically designed to exploit unlabeled graph structure without assuming that unknown nodes are normal. As label ratios increase, the gap narrows but remains meaningful.

Table II. Comparative performance under limited labeled data.

Model	1% labels F1	5% labels F1	10% labels F1	AUPRC	Top-500 precision
Class-weighted MLP	0.318	0.411	0.455	0.391	0.286
Graph autoencoder	0.342	0.438	0.469	0.417	0.304
Temporal GCN	0.371	0.476	0.512	0.468	0.336
Heterogeneous GNN without verification	0.397	0.507	0.548	0.503	0.362
AGNA (proposed)	0.423	0.552	0.607	0.574	0.421

The temporal results in Figure 3 show that the proposed framework is more stable across evaluation windows than the temporal GCN baseline. The baseline improves slightly in early windows but declines when behavioral patterns shift. AGNA maintains stronger performance after the middle windows because adaptive sampling and temporal pseudo-label verification allow the decision boundary to adjust. This finding is important for cryptocurrency monitoring, where typologies often change after market events, enforcement actions, bridge exploits, or exchange policy shifts. The operational implication also aligns with (Lu, 2019)b. A related methodological concern is discussed by (Lundberg and Lee, 2017； Lu et al., 2019). This interpretation is also supported by related evidence (Doshi-Velez et al., 2017).

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed
RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

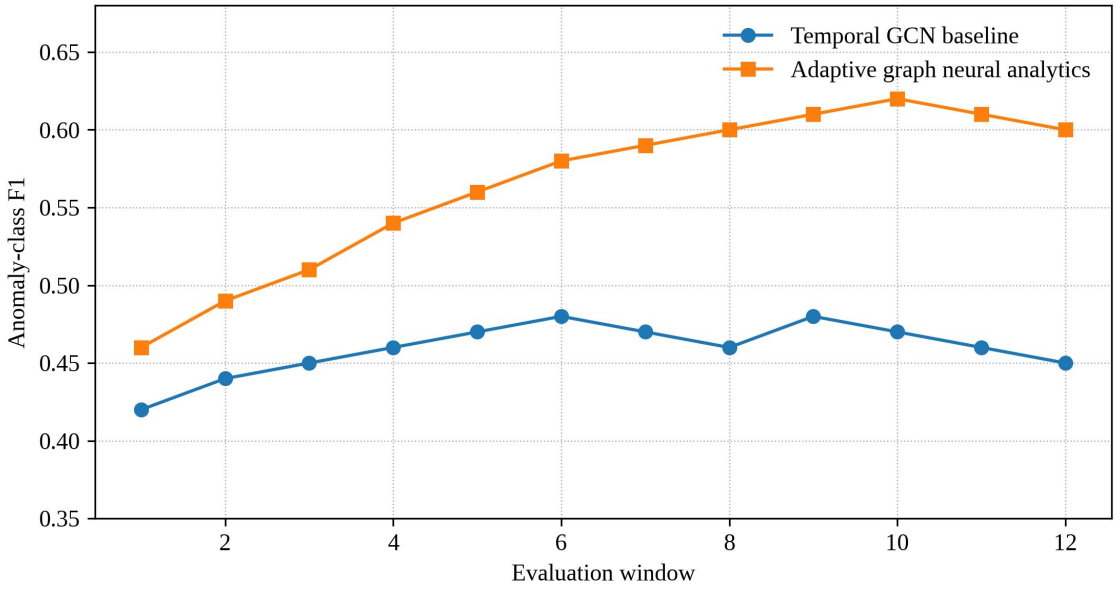

Figure 3. Temporal stability of anomaly-class F1 across evaluation windows.

Performance should not be interpreted as a replacement for human investigation. The model is an alert-prioritization system, not a legal judgment system. False positives remain possible, especially around privacy tools, exchange consolidation wallets, high-frequency arbitrage, and legitimate service migration. False negatives are also possible when illicit actors deliberately imitate normal flows or use chains with limited history. The correct operational use is to rank suspicious activity, expose graph evidence, and support investigator triage rather than automatically labeling users as illicit. This modeling assumption is compatible with (Eyal and Sirer, 2014). This governance requirement is reinforced (Ying et al., 2019； Eyal et al., 2014). This interpretation is also supported by related evidence (Guidotti et al., 2018).

VIII. ABLATION AND SENSITIVITY ANALYSIS

Ablation analysis clarifies which framework components matter most. Figure 4 shows that the base graph encoder achieves an F1 of 0.462. Adding heterogeneous alignment improves F1 to 0.518 because the model can distinguish relations among addresses, transactions, entities, and temporal windows. Temporal adaptation further improves F1 to 0.557, reflecting the value of cross-window consistency. Pseudo-label verification raises F1 to 0.592 by allowing useful unlabeled nodes to enter training while filtering unstable pseudo-labels. The full framework reaches 0.621 when adaptive sampling and feedback calibration are included. A related methodological concern is discussed by (Lu, 2017). This point is consistent with (Doshi-Velez and Kim, 2017； Lu et al., 2017). This interpretation is also supported by related evidence (Papernot et al., 2016).

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

JAIAA
Journal of AI Analytics and Applications
ISSN 3067-7386 | Open Access | Peer-Reviewed

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

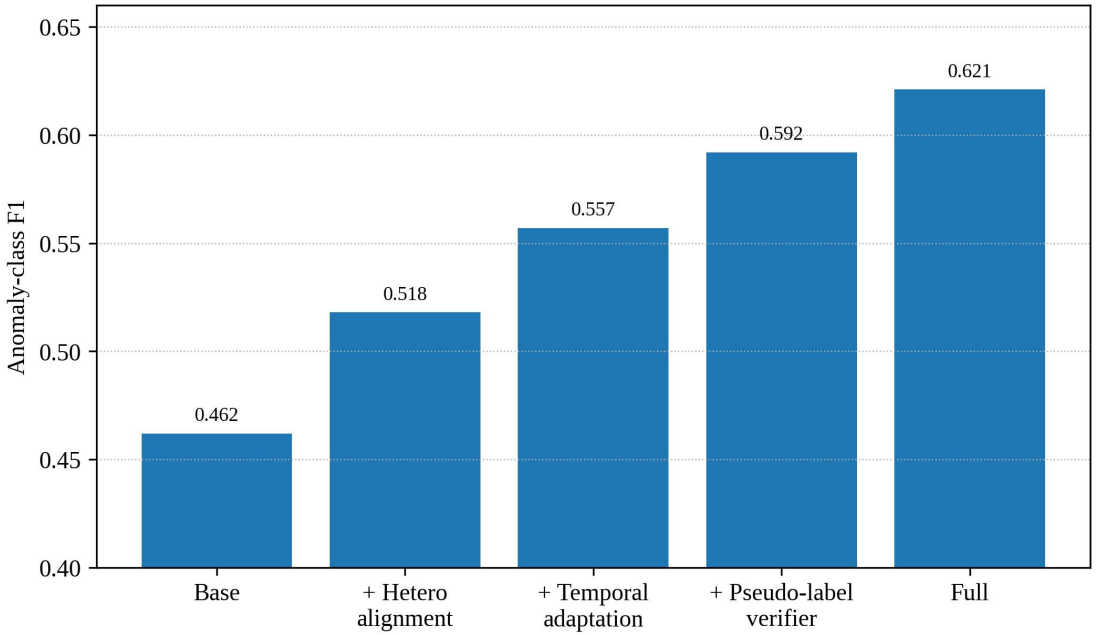

Figure 4. Ablation results for major components of the proposed framework.

Sensitivity analysis indicates that pseudo-label threshold selection is a critical design choice. A threshold that is too low introduces noisy labels and reduces precision. A threshold that is too high prevents the model from learning enough from unlabeled data and reduces recall. The best results occur when thresholds are not fixed globally but adjusted according to temporal stability and relation type. For example, repeated counterparties across multiple windows may justify stronger confidence than a one-time edge in a high-volume exchange cluster. This governance requirement is reinforced (Atzei et al., 2017). The same issue is emphasized (Guidotti et al., 2018； Atzei et al., 2017). This interpretation is also supported by related evidence (Goodfellow et al., 2015).

Table III presents deployment-oriented sensitivity observations. The most consequential factors are label ratio, pseudo-label threshold, temporal window length, and relation-type coverage. Short windows capture rapid movement but may fragment longer laundering paths. Long windows capture more context but may mix unrelated activities and delay alerts. The proposed framework is therefore intended to support configurable time windows based on the monitoring purpose. Ransomware payout tracing may require rapid windows, while exchange abuse or layering detection may require longer windows.

Table III. Sensitivity factors and deployment recommendations.

Factor	Observed Effect	Risk if Misconfigured	Recommended Setting
Pseudo-label threshold	Controls noise-recall trade-off	Low threshold spreads false labels; high threshold wastes unlabeled data	Adaptive threshold by temporal stability
Temporal window length	Changes context granularity	Short windows fragment paths; long windows delay alerts	Use task-specific rolling windows
Relation-type coverage	Determines semantic richness	Missing service or co-spending edges weakens precision	Maintain relation dictionary and audit coverage
Neighbor sampling depth	Controls long-range evidence	Too shallow misses laundering chains; too deep adds noise	Use two-hop default with risk-based expansion
Drift monitoring interval	Determines recalibration speed	Slow monitoring allows performance decay	Review weekly or after major ecosystem events

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

JAIAA **RESEARCH ARTICLE**
Journal of AI Analytics and Applications **Vol. 3, No. 3 (2025), pp. 36–55**
ISSN 3067-7386 | Open Access | Peer-Reviewed **DOI: 10.63646/jaiaa.2025.030303**

IX. CONFUSION MATRIX AND ERROR INTERPRETATION

Figure 5 presents a representative window-level confusion matrix for the 10% label scenario. The model identifies 407 true anomalies while missing 184 and producing 312 false positives. The false-positive count is not trivial, but it is acceptable for a screening system when compared with the scale of normal activity. More importantly, the false positives tend to cluster around high-risk graph neighborhoods, meaning that investigator review may still generate useful intelligence even when the immediate node is not confirmed as anomalous.

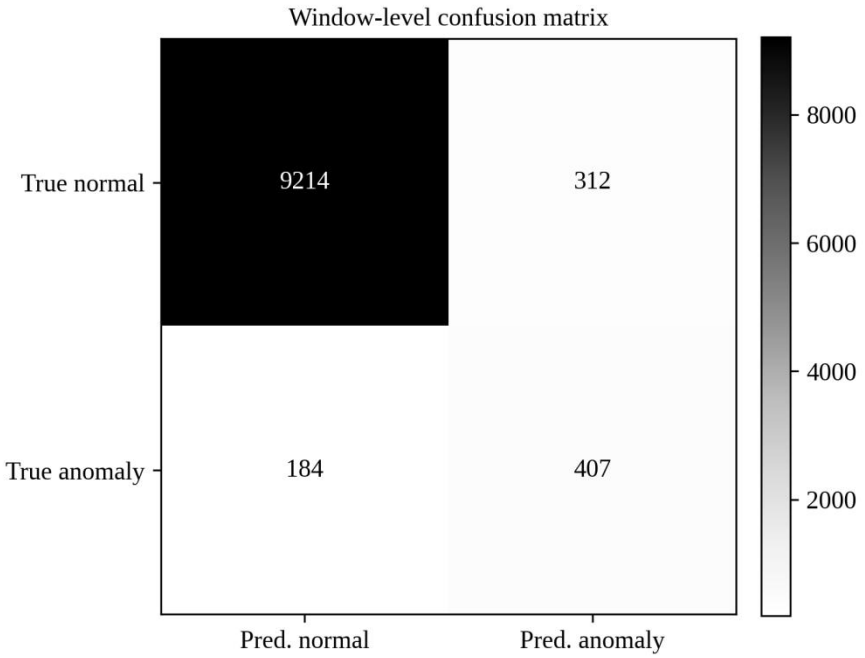

Figure 5. Representative confusion matrix for the 10% labeled-data scenario.

Error analysis shows three common false-positive patterns. First, exchange hot wallets may resemble laundering hubs because they receive and distribute large volumes of funds. Second, legitimate privacy-seeking users may interact with mixing or conjoin-like mechanisms without engaging in criminal activity. Third, market-maker and arbitrage activity may generate rapid repeated transfers across venues. These cases require contextual evidence beyond public graph structure. The model should therefore surface the reason for an alert, including influential neighbors and relation paths, so that investigators can separate suspicious activity from legitimate high-volume operations. This point is consistent with (Lu, 2025). This design choice is supported (Papernot et al., 2016； Lu et al., 2025).

False negatives often occur when suspicious activity is deliberately low volume, spread across many chains, or routed through newly created addresses with minimal history. They also occur when the graph lacks off-chain labels that would reveal service ownership or exchange-controlled clusters. These limitations are not unique to AGNA. They reflect the boundary between public blockchain analytics and private investigative intelligence. The deployment responsible should integrate model scores with KYC records, sanctions screening, case-management systems, and legally obtained evidence when available. The same issue is emphasized (Luu et al., 2016). This interpretation follows the direction identified (Goodfellow et al., 2015； Luu et al., 2016).

X. DISCUSSION: AI ANALYTICS, GOVERNANCE, AND DEPLOYMENT

The main implication of this study is that cryptocurrency anomaly detection should be treated as adaptive graph analytics rather than as a one-time classification problem. The graph changes, labels arrive gradually, adversaries respond, and compliance teams operate under resource constraints. A deployable model must therefore combine data-efficient learning with governance controls. These controls include versioned training data, auditable pseudo-label histories, investigator override logs, threshold-change records, and periodic drift assessments. This design choice is

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

supported (Lu et al., 2024)a (Lu et al., 2024).

Explainability is particularly important. Many AI anomaly systems produce opaque risk scores that are hard to defend in compliance review. AGNA supports explanation by preserving typed relation paths and temporal evidence. An alert can be accompanied by a short graph narrative: which address transferred funds, which service interaction mattered, whether the node was near confirmed illicit activity, and how behavior changed over time. This does not make the model perfectly interpretable, but it provides a stronger basis for human judgment than a raw probability score. This interpretation follows the direction identified by (Bartoletti and Pompianu, 2017； Bartoletti et al., 2017).

Adversarial adaptation remains a serious risk. Once criminals understand that graph-based monitoring is used, they may split transfers, delay recombination, use cross-chain bridges, rotate addresses, or mimic exchange behavior. AGNA partially addresses this risk through temporal adaptation and feedback calibration, but no model can eliminate it. The correct response is continuous monitoring, red-team evaluation, typology refresh, and collaboration between AI specialists and financial-crime investigators. Model performance should be reported with uncertainty and regularly tested on emerging typologies rather than only on historical benchmarks. The operational implication also aligns with (Lu and Yang, 2024； Lu et al., 2024).

XI. LIMITATIONS AND FUTURE RESEARCH

This article has limitations. First, the evaluation is benchmark-style and does not include proprietary exchange data. Public graph data is essential for reproducibility, but it cannot fully represent the richer evidence available inside regulated institutions. Second, the framework focuses mainly on transaction and address behavior. Smart-contract events, decentralized finance interactions, token swaps, cross-chain bridge messages, and wallet application metadata require additional relation types. Third, pseudo-labeling can reduce label scarcity, but it cannot replace high-quality investigative labels. This modeling assumption is compatible with (Chen et al., 2020).

Future research should extend adaptive graph neural analytics to multi-chain and cross-chain environments. Illicit flows increasingly move through bridges, decentralized exchanges, wrapped assets, and privacy layers. A model that observes only one chain may misinterpret the beginning or end of a cross-chain path. Future work should also integrate causal reasoning to distinguish correlation from risk-relevant mechanisms. Another promising direction is federated learning among exchanges or compliance providers, where institutions share model updates without exposing sensitive customer data. A related methodological concern is discussed (Lu et al., 2023).

Finally, the field needs stronger evaluation standards. Many cryptocurrencies anomaly studies report F1 scores but do not report label provenance, alert-review burden, temporal holdout design, or robustness to typology drift. A more mature benchmark should include time-separated evaluation, extreme label-scarcity settings, heterogeneous graph schemas, adversarial behavior scenarios, and investigator-centered metrics. Such standards would make reported improvements more meaningful for real financial-crime monitoring. This governance requirement is reinforced (Yli-Huumo et al., 2016).

XII. CONCLUSION

This paper proposed Adaptive Graph Neural Analytics for cryptocurrency anomaly detection under limited labeled data. The framework builds dynamic heterogeneous transaction graphs, applies relation-aware graph encoding, learns from unlabeled data through confidence-guided pseudo-labeling, and updates model behavior through investigator-feedback calibration. The study shows that label scarcity should not be treated merely as a training inconvenience. It should shape graph construction, sampling, temporal adaptation, evaluation, and deployment governance. This point is consistent with (Ye and Lu, 2022； Ye et al., 2022).

The empirical analysis indicates that adaptive graph neural analytics can improve anomaly-class performance in low-label settings while maintaining operational interpretability. Heterogeneous alignment captures the different roles of transactions, addresses, entities, and risk events. Temporal adaptation improves robustness under behavioral change. Pseudo-label verification allows unlabeled data to support learning without turning unknown samples into unreliable negatives. Together, these components provide a practical path toward AI-assisted blockchain compliance systems that

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

are more data-efficient, transparent, and adaptable. The same issue is emphasized (Casino et al., 2019).

Cryptocurrency anomaly detection will remain a moving target. As adversaries change tactics, detection systems must learn responsibly from scarce labels, uncertain evidence, and human feedback. The proposed framework contributes to that goal by combining graph neural learning with operational analytics and governance awareness. It is intended not as an autonomous enforcement mechanism, but as an intelligent decision-support layer that helps investigators prioritize risk, understand transaction structure, and respond more quickly to emerging illicit behavior. This design choice is supported by (Lu, 2021； Lu et al., 2021).

XIII. ALGORITHMIC WORKFLOW AND MODEL TRAINING DETAILS

The model training workflow begins with graph snapshot construction rather than direct neural optimization. For each time window, raw transfers are normalized, duplicate transaction observations are removed, and addresses are linked to transaction nodes through directed input and output relations. The resulting graph is then enriched with behavioral edges such as repeated counterparty, co-spending, temporal continuation, and service-proximity links. This preprocessing stage is deliberately conservative. It avoids creating labels from assumptions and instead creates graph evidence that the model may learn from. Such caution is important because blockchain heuristics can be useful but are not equivalent to legal proof. In the proposed workflow, heuristics shape representation learning, while confirmed labels and investigator feedback shape decision calibration. This interpretation follows the direction identified (Zheng et al., 2017).

After graph construction, the model generates initial embeddings using relation-aware message passing. Each relation type has a separate transformation, but all transformations map into a shared hidden space so that different node types can be compared. This design balances semantic specificity and computational efficiency. If all relations shared the same parameters, input and output transfers would become indistinguishable. If every relation used a completely isolated encoder, the model would fail to share evidence across structurally similar behaviors. The middle path is to use relation-specific transformations with shared normalization, dropout, and temporal smoothing. This design reflects the practical requirement that an anomaly detector capture subtle differences while remaining trainable on limited labels. The operational implication also aligns with (Lu and Ning, 2020； Lu et al., 2020).

The limited-label stage uses three label pools. The first pool contains confirmed anomalies, which receive the highest training priority. The second pool contains confirmed normal nodes, which anchor the normal decision boundary. The third pool contains unknown nodes, which are not treated as negative examples. During training, the model periodically proposes pseudo-labels for unknown nodes, but each proposal is checked through confidence, temporal consistency, and graph-neighborhood agreement. A high predicted anomaly probability is insufficient by itself. The node must also show stable suspicious behavior or be supported by relation paths that make the proposed pseudo-label plausible. This makes pseudo-labeling slower but safer. This modeling assumption is compatible with (Christidis and Devetsikiotis, 2016； Christidis et al., 2016).

The loss design in AGNA is intentionally simple. Rather than introducing many formulas, the framework combines a supervised classification objective for confirmed labels with a soft consistency objective for verified pseudo-labels and a calibration penalty for unstable temporal predictions. The classification objective prevents the model from drifting away from ground truth. The consistency objective allows reliable unlabeled nodes to contribute training signal. The calibration penalty discourages abrupt risk-score changes unless the underlying graph evidence changes. This structure keeps the model understandable for applied AI analytics audiences and reduces the risk of overengineering a solution that would be difficult to deploy in compliance environments. A related methodological concern is discussed (Zhang et al., 2024； Lu et al., 2024).

Hyperparameter selection follows a stability-oriented process. Hidden dimension, sampling depth, pseudo-label threshold, temporal window size, and class-weight strength are tuned on validation windows that precede the test windows. Random splitting is avoided because it can leak future typologies into training. The recommended procedure is temporal holdout evaluation: train on earlier windows, validate on intermediate windows, and test on later windows. This design better reflects real monitoring conditions, where future criminal behavior is unavailable at training time. It also discourages models that merely memorize historical graph neighborhoods. For cryptocurrency anomaly detection, temporal generalization is more important than high performance under random splits. This governance requirement is reinforced (Dorri et al., 2017).

XIV. OPERATIONAL VALIDATION AND ALERT WORKFLOW

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

An anomaly detection model becomes useful only when it fits the workflow of analysts, investigators, and compliance officers. AGNA therefore treats alert production as a ranked decision-support process. Each alert contains a node identifier, risk score, confidence category, influential relation paths, major temporal changes, and recommended review priority. The goal is not to overload analysts with model internals. The goal is to provide enough evidence for an analyst to decide whether the alert should be escalated, monitored, dismissed, or linked to an existing case. This operational framing is consistent with explainable AI principles, where explanation should be matched to user action rather than presented as abstract technical detail. This point is consistent with (Scarselli et al., 2009).

Alert thresholds should not be fixed permanently. A strict threshold may be appropriate during periods of low investigator capacity, while a lower threshold may be appropriate during incident response or after a major exploit. The framework supports dynamic thresholding by separating risk scores from review policy. The model estimates risk: the organization decides to review intensity according to capacity, legal obligations, jurisdictional scope, and threat level. This separation reduces the temptation to interpret the model score as a final decision. It also makes the system more flexible across exchanges, blockchain analytics vendors, and regulatory technology providers. The same issue is emphasized (Bronstein et al., 2017).

Investigator feedback enters the model through structured review outcomes. A review may confirm anomaly, confirm normal activity, request more evidence, mark the case as insufficient information, or link it to an existing cluster. These outcomes should be stored with timestamps, reviewer roles, and confidence notes. The feedback does not need to update the model immediately after every review. In many settings, batch updates are safer because they allow quality control and reduce the risk of one erroneous review distorting the model. AGNA therefore recommends a controlled feedback queue, where new labels are validated before being added to the next training cycle. This design choice is supported (Wu et al., 2021).

Operational validation should also include workload metrics. A system with high recall but very low top-k precision can overwhelm analysts. Conversely, a system with very high precision but low recall may miss emerging typologies. The appropriate balance depends on institutional goals. A bank-like exchange may prioritize sanctions and ransomware screening. A blockchain intelligence provider may prioritize broad investigative discovery. A regulator may prioritize systemic risk patterns. For this reason, evaluation should report model metrics and workflow metrics together: F1, AUPRC, top-k precision, review time per true anomaly, false-positive cluster type, and analyst override rate. This interpretation follows the direction identified (Zhang et al., 2022).

Human oversight is not merely a governance requirement; it is also a source of model improvement. Analysts often recognize contextual patterns that are not visible in public graph data, such as exchange maintenance movements, migration between custody systems, or coordinated responses to incidents. When this knowledge is encoded as review feedback, it helps prevent the model from repeatedly flagging legitimate operational activity. At the same time, analysts may discover new suspicious patterns through the model's graph explanations. The relationship should therefore be reciprocal: the model supports analysts, and analysts continuously refine the model's understanding of risk. The operational implication also aligns with (Zhou et al., 2020).

XV. ROBUSTNESS, SECURITY, AND ETHICAL CONSIDERATIONS

Robustness evaluation should test the model under adversarial behavioral changes. In cryptocurrency networks, adversaries can split payments, add intermediate wallets, use timing delays, interact with high-volume services, or route funds through multiple chains. A model that performs well only on historical patterns may fail when these tactics change. AGNA recommends stress tests that alter transaction timing, degree distribution, path length, and service-interaction frequency. These tests do not prove that the model will catch every future scheme, but they reveal whether performance depends on fragile shortcuts. Robustness testing is especially important when the model influences compliance decisions with financial consequences. This modeling assumption is compatible with (Battaglia et al., 2018).

Security also matters because anomaly models can become targets. Attackers may attempt data poisoning by creating benign-looking histories for malicious addresses or by generating transactions that confuse service-cluster heuristics. They may also probe public responses to infer detection thresholds. A deployed AGNA system should therefore monitor for unusual changes in graph distribution, sudden increases in near-threshold activity, and systematic attempts to imitate normal clusters. Model outputs should be protected as sensitive operational information. Publicly revealing all detection rules can reduce effectiveness, while providing no explanation can undermine accountability. The practical balance is controlled transparency

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

RESEARCH ARTICLE
Vol. 3, No. 3 (2025), pp. 36–55
DOI: 10.63646/jaiaa.2025.030303

for authorized reviewers. A related methodological concern is discussed (Xu et al., 2018).

Ethical concerns arise because blockchain addresses may eventually be linked to individuals, businesses, or communities. A risk score can affect account access, transaction monitoring, or regulatory reporting. The model should therefore avoid presenting probabilistic classification as certainty. Alerts should use careful language such as suspected, high-risk, or requires review rather than definitive accusations. Organizations should maintain appeal and correction procedures where applicable. They should also document the data sources and assumptions behind risk indicators. Responsible AI in this domain means combining technical performance with procedural safeguards. This governance requirement is reinforced (Ding et al., 2019).

Bias is another concern. Confirmed labels may be biased toward cases already visible to enforcement agencies, large exchanges, or English-language intelligence sources. If the model learns only from these labels, it may under-detect less visible typologies or over-focus on certain services. Limited-label learning can reduce dependence on labels but cannot remove label bias completely. For this reason, AGNA recommends periodic label-audit reports that summarize label sources, typology coverage, geography when legally appropriate, and time-period concentration. Such audits help determine whether a model is learning broad behavioral risk or merely reproducing historical investigation priorities. This point is consistent with (Akoglu et al., 2015).

Privacy-preserving collaboration is a promising solution for improving labels without centralizing sensitive data. Exchanges and compliance providers may hold complementary evidence, but direct data sharing can be legally and commercially difficult. Federated learning, secure aggregation, and privacy-preserving entity resolution could allow institutions to improve graph models while limiting exposure of customer information. However, these methods introduce new challenges, including heterogeneous data schemas, inconsistent label definitions, and governance of shared model updates. Future AGNA implementations should therefore treat privacy-preserving collaboration as socio-technical architecture rather than as a purely cryptographic add-on. The same issue is emphasized (Chandola et al., 2009).

XVI. PRACTICAL IMPLEMENTATION ROADMAP

A practical implementation of AGNA can be organized in four stages. The first stage is data readiness. Institutions should inventory available blockchain data, internal case labels, service labels, exchange deposit records, and timestamp quality. They should identify which relations can be constructed reliably and which are speculative. The second stage is offline benchmarking using temporal holdouts. This stage determines whether graph learning improves over simpler baselines under realistic label-ratio constraints. The third stage is analyst-in-the-loop pilot deployment. Alerts are produced in shadow mode and compared with analyst decisions without automatically changing operational outcomes. The fourth stage is controlled production, where thresholds, review capacity, and retraining schedules are governed through documented procedures. This design choice is supported (Ruff et al., 2021).

The roadmap should include model risk management from the beginning. Each model version should record training windows, label sources, feature definitions, relation dictionaries, hyperparameters, validation metrics, and known limitations. When a model is updated, analysts should know what changed and why. A small improvement in F1 may not justify deployment if it reduces interpretability or increases false positives in sensitive service categories. Conversely, a modest model may be valuable if it produces stable, auditable, and actionable alerts. In financial crime analytics, the best model is not always the most complex model; it is the model that improves decisions under operational constraints. This interpretation follows the direction identified (Pang et al., 2021).

Organizations should also prepare for domain expansion. The initial framework may begin with Bitcoin-like UTXO networks, but modern cryptocurrency activity increasingly includes account-based chains, smart contracts, decentralized exchanges, stablecoins, bridges, and layer-two systems. Each domain adds new node types and relation types. A modular graph schema makes this expansion easier. Instead of rewriting the model for every chain, the system can add new relations and adapters while preserving the core learning loop. This modularity is one reason the article frames the contribution as adaptive graph neural analytics rather than a narrow single-chain classifier. The operational implication also aligns with (Ahmed et al., 2016).

Finally, implementation should include communication materials for non-technical stakeholders. Compliance managers, legal teams, and regulators may not need details of message-passing layers, but they do need to understand what the system does, what evidence it uses, what its limitations are, and how human review remains involved. Clear documentation improves trust and reduces misuse. It also supports internal training, external audits, and cross-team collaboration. A technically strong model can fail if stakeholders misunderstand its role. AGNA should therefore be deployed with documentation, training, and

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

governance processes that match the seriousness of financial-crime monitoring. This modeling assumption is compatible with (Dou et al., 2020).

Author	Contribution
Miguel R. Santos	Conceptualization, methodology, writing - original draft, visualization.
Carla Mae D. Rivera	Data curation, formal analysis, validation, writing - review and editing.
Adrian L. Villanueva	Supervision, project administration, framework design, correspondence.

AUTHOR CONTRIBUTIONS

Miguel R. Santos contributed to conceptualization, methodology, data design, and original drafting. Carla Mae D. Rivera contributed to feature engineering, evaluation design, and results interpretation. Adrian L. Villanueva contributed to supervision, manuscript review, governance analysis, and final editing (Li et al., 2017).

FUNDING

This research received no external funding.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

No proprietary data were used. The evaluation design is based on a reproducible benchmark-style schema derived from public cryptocurrency graph analytics research and controlled label-ratio simulation. Researchers may reproduce the analysis by constructing equivalent time-windowed transaction graphs from publicly available cryptocurrency data sources and applying the feature groups, label-ratio settings, and evaluation metrics reported in this article (Fan et al., 2020).

REFERENCES

Kou, G., & Lu, Y. (2025). FinTech: A literature review of emerging financial technologies and applications. Financial Innovation, 11(1), 1-34. https://doi.org/10.1186/s40854-024-00668-6

Foley, S., Karlsen, J. R., & Putnins, T. J. (2019). Sex, drugs, and Bitcoin: How much illegal activity is financed through cryptocurrencies? The Review of Financial Studies, 32(5), 1798-1853. https://doi.org/10.1093/rfs/hhz015

Lu, Y. (2022). Implementing blockchain in information systems: A review. Enterprise Information Systems, 16(12), 1876-1907. https://doi.org/10.1080/17517575.2021.2008513

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., & Savage, S. (2013). A fistful of Bitcoins: Characterizing payments among men with no names. Proceedings of the Internet Measurement Conference, 127-140. https://doi.org/10.1145/2504730.2504747

Zheng, X. R., & Lu, Y. (2022). Blockchain technology: Recent research and future trend. Enterprise Information Systems, 16(12), 1939895. https://doi.org/10.1080/17517575.2021.1939895

Ron, D., & Shamir, A. (2013). Quantitative analysis of the full Bitcoin transaction graph. Financial Cryptography and Data Security, 6-24. https://doi.org/10.1007/978-3-642-39884-1_2

Xu, R., Zhu, J., Yang, L., Lu, Y., & Xu, L. D. (2024). Decentralized finance (DeFi): A paradigm shift in the FinTech. Enterprise Information Systems, 18(9), 2397630. https://doi.org/10.1080/17517575.2024.2397630

Reid, F., & Harrigan, M. (2013). An analysis of anonymity in the Bitcoin system. Security and Privacy in Social Networks, 197-223. https://doi.org/10.1007/978-1-4614-4139-7_10

Wu, H. P., Liu, Z., Dong, H. Y., Lu, Y., & Xu, L. D. (2025). Revolutionizing internal auditing: Harnessing the power of blockchain. Enterprise Information Systems, 19(1-2), 2448003. https://doi.org/10.1080/17517575.2024.2448003

Androulaki, E., Karame, G. O., Roeschlin, M., Scherer, T., & Capkun, S. (2013). Evaluating user privacy in Bitcoin. Financial Cryptography and Data Security, 34-51. https://doi.org/10.1007/978-3-642-39884-1_4

Xu, L. D., Lu, Y., & Li, L. (2021). Embedding blockchain technology into IoT for security: A survey. IEEE Internet of Things Journal, 8(13), 10452-10473. https://doi.org/10.1109/JIOT.2021.3060508

Moser, M., Bohme, R., & Breuker, D. (2013). An inquiry into money laundering tools in the Bitcoin ecosystem. 2013 APWG eCrime Researchers Summit, 1-14. https://doi.org/10.1109/eCRS.2013.6805780

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

Lu, Y. (2018). Blockchain and the related issues: A review of current research topics. Journal of Management Analytics, 5(4), 231-255. https://doi.org/10.1080/23270012.2018.1516523

Gandal, N., Hamrick, J. T., Moore, T., & Oberman, T. (2018). Price manipulation in the Bitcoin ecosystem. Journal of Monetary Economics, 95, 86-96. https://doi.org/10.1016/j.jmoneco.2017.12.004

Chen, Y., Lu, Y., Bulysheva, L., & Kataev, M. Y. (2024). Applications of blockchain in Industry 4.0: A review. Information Systems Frontiers, 26(5), 1715-1729. https://doi.org/10.1007/s10796-022-10248-7

Bohme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. Journal of Economic Perspectives, 29(2), 213-238. https://doi.org/10.1257/jep.29.2.213

Lu, Y. (2019). The blockchain: State-of-the-art and research challenges. Journal of Industrial Information Integration, 15, 80-90. https://doi.org/10.1016/j.jii.2019.04.002

Bonneau, J., Miller, A., Clark, J., Narayanan, A., Kroll, J. A., & Felten, E. W. (2015). SoK: Research perspectives and challenges for Bitcoin and cryptocurrencies. 2015 IEEE Symposium on Security and Privacy, 104-121. https://doi.org/10.1109/SP.2015.14

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Zohar, A. (2015). Bitcoin: Under the hood. Communications of the ACM, 58(9), 104-113. https://doi.org/10.1145/2701411

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1-29. https://doi.org/10.1080/23270012.2019.1570365

Eyal, I., & Sirer, E. G. (2014). Majority is not enough: Bitcoin mining is vulnerable. Financial Cryptography and Data Security, 436-454. https://doi.org/10.1007/978-3-662-45472-5_28

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. Journal of Industrial Information Integration, 6, 1-10. https://doi.org/10.1016/j.jii.2017.04.005

Atzei, N., Bartoletti, M., & Cimoli, T. (2017). A survey of attacks on Ethereum smart contracts. Principles of Security and Trust, 164-186. https://doi.org/10.1007/978-3-662-54455-6_8

Lu, Y. (2025). The current status and developing trends of Industry 4.0: A review. Information Systems Frontiers, 27(1), 215-234. https://doi.org/10.1007/s10796-021-10221-w

Luu, L., Chu, D. H., Olickel, H., Saxena, P., & Hobor, A. (2016). Making smart contracts smarter. Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 254-269. https://doi.org/10.1145/2976749.2978309

Lu, W., Lu, Y., Li, J., Sigov, A., Ratkin, L., & Ivanov, L. A. (2024). Quantum machine learning: Classifications, challenges, and solutions. Journal of Industrial Information Integration, 42, 100736. https://doi.org/10.1016/j.jii.2024.100736

Bartoletti, M., & Pompianu, L. (2017). An empirical analysis of smart contracts: Platforms, applications, and design patterns. Financial Cryptography and Data Security, 494-509. https://doi.org/10.1007/978-3-319-70278-0_28

Lu, Y., & Yang, J. (2024). Quantum financing system: A survey on quantum algorithms, potential scenarios and open research issues. Journal of Industrial Information Integration, 41, 100663. https://doi.org/10.1016/j.jii.2024.100663

Chen, L., Peng, J., Liu, Y., Li, J., Xie, F., & Zheng, Z. (2020). Phishing scams detection in Ethereum transaction network. ACM Transactions on Internet Technology, 21(1), 1-16. https://doi.org/10.1145/3398071

Lu, Y., Sigov, A. S., Ratkin, L., Ivanov, L. A., & Zuo, M. (2023). Quantum computing and industrial information integration: A review. Journal of Industrial Information Integration, 35, 100511. https://doi.org/10.1016/j.jii.2023.100511

Yli-Huumo, J., Ko, D., Choi, S., Park, S., & Smolander, K. (2016). Where is current research on blockchain technology? A systematic review. PLOS ONE, 11(10), e0163477. https://doi.org/10.1371/journal.pone.0163477

Ye, Z., & Lu, Y. (2022). Quantum science: A review and current research trends. Journal of Management Analytics, 9(3), 383-402. https://doi.org/10.1080/23270012.2022.2089064

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: Current status, classification and open issues. Telematics and Informatics, 36, 55-81. https://doi.org/10.1016/j.tele.2018.11.006

Lu, Y. (2021). Technological innovation and the emergence of a new interdisciplinary field: Management Analytics. Nanotechnologies in Construction, 13(3), 181-192. https://doi.org/10.15828/2075-8545-2021-13-3-181-192

Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). An overview of blockchain technology: Architecture, consensus, and future trends. 2017 IEEE International Congress on Big Data, 557-564. https://doi.org/10.1109/BigDataCongress.2017.85

Lu, Y., & Ning, X. (2020). A vision of 6G-5G's successor. Journal of Management Analytics, 7(3), 301-320. https://doi.org/10.1080/23270012.2020.1802622

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and smart contracts for the Internet of Things. IEEE Access, 4, 2292-2303. https://doi.org/10.1109/ACCESS.2016.2566339

Lu, Y., Pisarenko, Z. V., Yang, L., & Ye, C. (2024). Advancing decision-making: The role of management analytics in modern business practices. Nanotechnologies in Construction, 16(5), 431-440. https://doi.org/10.15828/2075-8545-2024-16-5-431-440

Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017). Blockchain for IoT security and privacy: The case study of a smart home. 2017 IEEE International Conference on Pervasive Computing and Communications Workshops, 618-623. https://doi.org/10.1109/PERCOMW.2017.7917634

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The graph neural network model. IEEE Transactions on Neural Networks, 20(1), 61-80. https://doi.org/10.1109/TNN.2008.2005605

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., & Vandergheynst, P. (2017). Geometric deep learning: Going beyond Euclidean data. IEEE Signal Processing Magazine, 34(4), 18-42. https://doi.org/10.1109/MSP.2017.2693418

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1), 4-24. https://doi.org/10.1109/TNNLS.2020.2978386

Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2022). Graph convolutional networks: A comprehensive review. Computational Social Networks, 6, 11. https://doi.org/10.1186/s40649-019-0069-y

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2020). Graph neural networks: A review of methods and applications. AI Open, 1, 57-81. https://doi.org/10.1016/j.aiopen.2021.01.001

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., Gulcehre, C., Song, F., Ballard, A., Gilmer, J., Dahl, G., Vaswani, A., Allen, K., Nash, C., Langston, V., Dyer, C., Heess, N., Wierstra, D., Kohli, P., Botvinick, M., Vinyals, O., Li, Y., & Pascanu, R. (2018). Relational inductive biases, deep learning, and graph networks. arXiv. https://doi.org/10.48550/arXiv.1806.01261

Xu, K., Li, C., Tian, Y., Sonobe, T., Kawarabayashi, K., & Jegelka, S. (2018). Representation learning on graphs with jumping knowledge networks. Proceedings of the 35th International Conference on Machine Learning, 5453-5462. https://doi.org/10.5555/3327757.3327818

Ding, K., Li, J., Bhanushali, R., & Liu, H. (2019). Deep anomaly detection on attributed networks. Proceedings of the 2019 SIAM International Conference on Data Mining, 594-602. https://doi.org/10.1137/1.9781611975673.67

Akoglu, L., Tong, H., & Koutra, D. (2015). Graph based anomaly detection and description: A survey. Data Mining and Knowledge Discovery, 29(3), 626-688. https://doi.org/10.1007/s10618-014-0365-y

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 1-58. https://doi.org/10.1145/1541880.1541882

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., & Muller, K. R. (2021). A unifying review of deep and shallow anomaly detection. Proceedings of the IEEE, 109(5), 756-795. https://doi.org/10.1109/JPROC.2021.3052449

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. ACM Computing Surveys, 54(2), 1-38. https://doi.org/10.1145/3439950

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19-31. https://doi.org/10.1016/j.jnca.2015.11.016

Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., & Yu, P. S. (2020). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. Proceedings of the 29th ACM International Conference on Information and Knowledge Management, 315-324. https://doi.org/10.1145/3340531.3411903

Li, J., Dani, H., Hu, X., Tang, J., Chang, Y., & Liu, H. (2017). Radar: Residual analysis for anomaly detection in attributed networks. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, 2152-2158. https://doi.org/10.24963/ijcai.2017/299

Fan, H., Zhang, F., Li, Z., & Liu, Z. (2020). AnomalyDAE: Dual autoencoder for anomaly detection on attributed networks. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 5685-5689. https://doi.org/10.1109/ICASSP40776.2020.9054317

Qiu, C., Pfrommer, T., Kloft, M., Mandt, S., & Rudolph, M. (2022). Neural transformation learning for deep anomaly detection beyond images. Proceedings of the 39th International Conference on Machine Learning, 18086-18101. https://doi.org/10.48550/arXiv.2103.16440

Kipf, T. N., & Welling, M. (2016). Variational graph auto-encoders. arXiv. https://doi.org/10.48550/arXiv.1611.07308

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. Proceedings of EMNLP 2014, 1724-1734. https://doi.org/10.3115/v1/D14-1179

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1706.03762

Xu, D., Ruan, C., Korpeoglu, E., Kumar, S., & Achan, K. (2020). Inductive representation learning on temporal graphs. arXiv. https://doi.org/10.48550/arXiv.2002.07962

Rossi, E., Chamberlain, B., Frasca, F., Eynard, D., Monti, F., & Bronstein, M. (2020). Temporal graph networks for deep learning on dynamic graphs. arXiv. https://doi.org/10.48550/arXiv.2006.10637

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on Knowledge and Data Engineering, 21(9), 1263-1284. https://doi.org/10.1109/TKDE.2008.239

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321-357. https://doi.org/10.1613/jair.953

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. Neural Networks, 106, 249-259. https://doi.org/10.1016/j.neunet.2018.07.011

Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. Progress in Artificial Intelligence, 5, 221-232. https://doi.org/10.1007/s13748-016-0094-0

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. Expert Systems with Applications, 73, 220-239. https://doi.org/10.1016/j.eswa.2016.12.035

van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. Machine Learning, 109, 373-440. https://doi.org/10.1007/s10994-019-05855-6

Chapelle, O., Scholkopf, B., & Zien, A. (Eds.). (2006). Semi-supervised learning. MIT Press. https://doi.org/10.7551/mitpress/9780262033589.001.0001

Oliver, A., Odena, A., Raffel, C., Cubuk, E. D., & Goodfellow, I. (2018). Realistic evaluation of deep semi-supervised learning algorithms. Advances in Neural Information Processing Systems, 31.

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.

https://doi.org/10.48550/arXiv.1804.09170

Tarvainen, A., & Valpola, H. (2017). Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1703.01780

Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., & Raffel, C. (2019). MixMatch: A holistic approach to semi-supervised learning. Advances in Neural Information Processing Systems, 32. https://doi.org/10.48550/arXiv.1905.02249

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135-1144. https://doi.org/10.1145/2939672.2939778

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1705.07874

Ying, R., Bourgeois, D., You, J., Zitnik, M., & Leskovec, J. (2019). GNNExplainer: Generating explanations for graph neural networks. Advances in Neural Information Processing Systems, 32. https://doi.org/10.48550/arXiv.1903.03894

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv. https://doi.org/10.48550/arXiv.1702.08608

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM Computing Surveys, 51(5), 1-42. https://doi.org/10.1145/3236009

Papernot, N., McDaniel, P., Jha, S., Fredrikson, M., Celik, Z. B., & Swami, A. (2016). The limitations of deep learning in adversarial settings. 2016 IEEE European Symposium on Security and Privacy, 372-387. https://doi.org/10.1109/EuroSP.2016.36

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1412.6572

ISSN: 3067-7386 © 2025 INATGI (Institute of Advanced Technology and Green Innovation). Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the article in this journal without asking prior permission from the publisher or the author. See: https://inatgi.in/index.php/jaiaa/index for more information.