ISSN - 2456-1142

International journal of humanities and information technology Volume 8, Issue 1, 2026

Graph Neural Network-Based Nationwide Healthcare Fraud Detection and Financial Risk Intelligence for Medicare and Medicaid Systems

Trang Huynh*

Independent Researcher.

Abstract

Healthcare fraud, waste, and abuse remain major financial and operational threats to Medicare and Medicaid systems, where fraudulent claims, provider collusion, upcoding, phantom billing, and abnormal referral patterns can generate substantial public expenditure losses. Traditional fraud detection methods, including rule-based systems, statistical models, and conventional machine learning, often analyze claims as isolated records and may fail to capture the complex relationships among providers, beneficiaries, procedures, diagnoses, pharmacies, facilities, and payment networks. This paper proposes a Graph Neural Network-based nationwide healthcare fraud detection and financial risk intelligence framework for Medicare and Medicaid systems. The proposed architecture models healthcare claims as a heterogeneous graph, enabling relational learning across multiple entities and supporting the identification of suspicious providers, anomalous claims, hidden fraud communities, and high-risk financial patterns. Building on graph anomaly detection, imbalanced learning, explainable artificial intelligence, and cost-sensitive fraud analytics, the framework integrates graph construction, GNN-based representation learning, fraud risk scoring, explainability, and audit-support dashboards. The paper contributes a scalable and policy-relevant model for improving payment integrity, strengthening public insurance oversight, prioritizing fraud investigations, and supporting data-driven financial risk intelligence across nationwide healthcare programs.

Keywords: Graph Neural Networks; Healthcare Fraud Detection; Medicare; Medicaid; Financial Risk Intelligence; Claims Analytics; Explainable AI.

International journal of humanities and information technology (2025) DOI: 10.21590/ijhit.08.02.03

Introduction

Background of Healthcare Fraud in Public Insurance Systems

Healthcare fraud, waste, and abuse remain major financial and administrative challenges within public insurance systems, particularly Medicare and Medicaid. These programs process large volumes of claims across diverse providers, beneficiaries, facilities, procedures, diagnoses, pharmacies, and payment channels. Because of this scale and complexity, fraudulent activities can remain hidden within normal healthcare transactions for long periods. Common fraud patterns include fraudulent billing, upcoding, phantom claims, unnecessary medical services, duplicate claims, provider collusion, abnormal referral networks, and manipulation of diagnosis or procedure codes for higher reimbursement. Such practices weaken public healthcare financing, reduce resources available for genuine patient care, and increase the burden on federal and state health agencies.

Fraud detection has long been studied through statistical, data mining, and machine learning approaches. Bolton and Hand (2002) emphasized that fraud detection requires identifying rare, abnormal, and deceptive patterns within large transactional datasets. In healthcare, this challenge becomes more difficult because fraudulent claims may appear clinically plausible when viewed individually. Joudaki et al. (2014) showed that data mining techniques have been widely applied to detect healthcare fraud and abuse, but healthcare fraud continues to evolve as providers and fraud groups adapt their behavior. Thornton et al. (2015) further categorized healthcare fraud into multiple forms, including

Corresponding Author: Trang Huynh, Independent Researcher, e-mail: huynhtrang085@gmail.com

How to cite this article: Huynh, Trang. (2026). Graph Neural Network-Based Nationwide Healthcare Fraud Detection and Financial Risk Intelligence for Medicare and Medicaid Systems. *International journal of humanities and information technology* 8(1), 19-29.

Source of support: Nil

Conflict of interest: None

service manipulation, billing irregularities, and provider-level abuse. In Medicare systems, Bauder and Khoshgoftaar (2017) demonstrated the relevance of machine learning for detecting suspicious Medicare providers, while du Preez et al. (2025) highlighted the growing importance of advanced machine learning methods for healthcare claims fraud detection.

Problem Statement

Despite progress in fraud analytics, conventional fraud detection systems remain limited. Many existing systems depend on rule-based filters, manual audits, statistical thresholds, or supervised machine learning models trained on isolated claim-level features. While these methods are useful for detecting known fraud patterns, they often struggle to identify complex, hidden, and coordinated fraud schemes. Healthcare fraud is rarely limited to a single claim. It may involve relationships among providers, beneficiaries, pharmacies, diagnosis codes, procedure codes, referral pathways, geographic clusters, and repeated payment behaviors. Therefore, a claim that appears normal in isolation may become suspicious when examined within a wider network of healthcare interactions.

Traditional models also face major challenges such as class imbalance, limited confirmed fraud labels, adaptive fraud behavior, high false-positive rates, and weak interpretability. Chandola et al. (2009) explained that anomaly detection is difficult when abnormal cases are rare, dynamic, and context-dependent. Akoglu et al. (2015) argued that graph-based anomaly detection is valuable because many abnormal behaviors are better identified through relationships, communities, and structural deviations. Johnson and Khoshgoftaar (2019) also showed that neural network methods can improve Medicare fraud detection, but conventional neural models may still fail to fully represent relational dependencies in claims networks. These limitations create the need for a nationwide fraud detection framework that can learn from both healthcare attributes and graph-based relationships.

Research Aim and Objectives

The aim of this paper is to propose a nationwide graph neural network-based framework for healthcare fraud detection and financial risk intelligence in Medicare and Medicaid systems. The proposed framework is designed to transform healthcare claims data into a heterogeneous graph structure and apply graph neural network learning to detect suspicious providers, claims, and fraud communities.

The specific objectives are to model Medicare and Medicaid claims as a heterogeneous healthcare graph, detect suspicious providers and coordinated fraud communities using GNN-based learning, integrate imbalance-aware and cost-sensitive fraud detection strategies, support explainable financial risk intelligence for auditors and regulators, and propose a scalable architecture suitable for nationwide healthcare fraud surveillance.

1.4 Main Contribution of the Paper

This paper contributes a scalable conceptual and technical framework that integrates graph neural networks, heterogeneous claims modeling, anomaly detection, explainable artificial intelligence, financial risk scoring, and public insurance oversight. Unlike conventional claim-level fraud detection systems, the proposed model emphasizes relational intelligence by capturing connections among healthcare actors and transactions. The framework therefore supports more effective fraud detection, audit prioritization, payment integrity monitoring, and financial risk intelligence across Medicare and Medicaid systems.

Literature Review and Theoretical Foundation

Traditional Statistical and Machine Learning Fraud Detection

Healthcare fraud detection initially relied on statistical rules, manual audits, and abnormal pattern identification. Early fraud analytics focused on identifying unusual claims, billing outliers, excessive service use, duplicate submissions, and abnormal provider reimbursement behavior. Bolton and Hand (2002) described statistical fraud detection as a process of identifying rare and suspicious patterns in large transaction datasets, while Ortega et al. (2006) showed that data mining could support medical claim fraud detection through structured claim analysis. However, traditional rule-based systems are often rigid and may fail when fraudsters change their behavior.

Machine learning improved fraud detection by enabling models to learn patterns from historical data. Supervised models can classify claims or providers as fraudulent or legitimate when labeled examples are available. Bauder and Khoshgoftaar (2017) applied machine learning methods to Medicare fraud detection, showing that predictive models can support automated fraud screening. Neural network approaches further improved pattern recognition by capturing nonlinear relationships in healthcare claims data (Johnson & Khoshgoftaar, 2019). Unsupervised anomaly detection is also important because confirmed fraud labels are limited. Chandola et al. (2009) emphasized that anomaly detection is suitable for identifying rare events that differ from normal behavior, while Naidoo and Marivate (2020) demonstrated the use of generative adversarial networks for detecting suspicious healthcare providers. Recent explainable models have also become important because fraud detection systems must justify why a provider or claim is considered high risk (Hancock et al., 2023). Despite these improvements, most traditional models analyze claims as isolated records and may miss hidden relationships among providers, beneficiaries, procedures, and facilities.

Table 1: Summary of Fraud Detection Approaches in Healthcare and Financial Risk Analytics

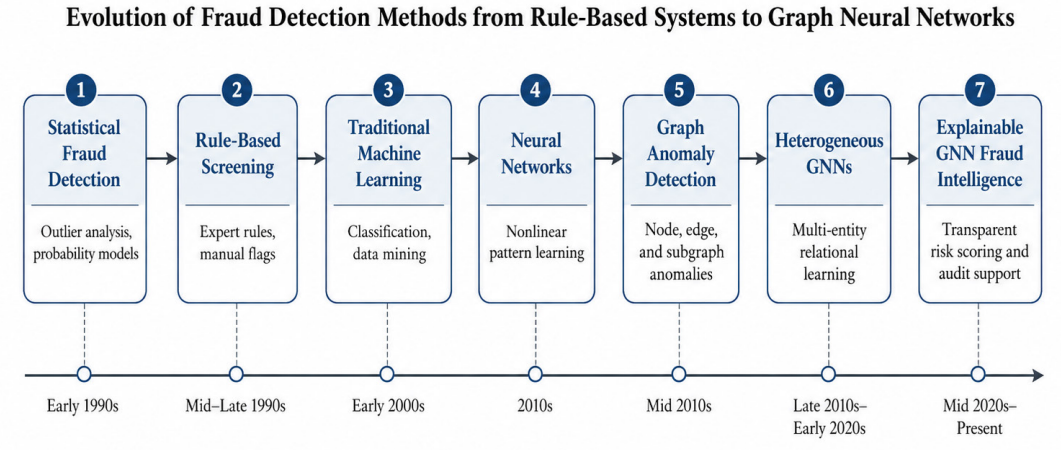
Approach	*Core technique*	*Strength*	*Limitation*	*Key references*
Rule-based systems	Predefined fraud rules	Easy to apply	Rigid and reactive	Bolton & Hand (2002)
Statistical models	Outlier and probability analysis	Detects abnormal claims	Weak for complex fraud networks	Bolton & Hand (2002); Ortega et al. (2006)
Supervised ML	Classification models	Learns from labeled fraud data	Needs reliable labels	Bauder & Khoshgoftaar (2017)
Neural networks	Nonlinear pattern learning	Captures complex claim patterns	Less interpretable	Johnson & Khoshgoftaar (2019)
Anomaly detection	Rare pattern identification	Useful with limited labels	May increase false positives	Chandola et al. (2009)
GAN-based detection	Synthetic anomaly learning	Supports unsupervised provider screening	Computationally demanding	Naidoo & Marivate (2020)
Explainable ML	Feature-based interpretation	Supports audit transparency	May not capture graph structure	Hancock et al. (2023)

Graph-Based Anomaly Detection

Healthcare fraud is often relational rather than isolated. Fraud may involve providers sharing addresses, beneficiaries receiving repeated services from suspicious networks, circular referral patterns, abnormal procedure clusters, or coordinated billing across facilities. Graph-based anomaly detection is useful because it represents entities as nodes and relationships as edges. Akoglu et al. (2015) explained that graph anomaly detection can identify suspicious nodes, edges, subgraphs, and communities. This is important for Medicare and Medicaid because fraud may appear as coordinated behavior across a network rather than a single abnormal claim. Wu et al. (2020) further noted that graph learning methods can model both attributes and structure, while Pereira and Murai (2021) showed that GNNs can be effective for fraud detection in networked data.

Graph Neural Networks for Fraud Detection

Graph neural networks provide a stronger technical foundation for fraud detection because they learn from both claim attributes and network relationships. Graph Convolutional Networks aggregate information from neighboring nodes (Kipf & Welling, 2016), while GraphSAGE supports inductive learning on large graphs and unseen nodes (Hamilton et al., 2017). Graph Attention Networks assign different weights to important relationships, making them useful for identifying influential fraud links (Veličković et al., 2017). More advanced healthcare fraud studies now use heterogeneous, hierarchical, and dynamic graph structures to model multiple entity types and evolving fraud behavior (Lu et al., 2023; Hong et al., 2024; Zhang et al., 2022). Fraud-specific GNN research also addresses camouflaged fraudsters and relational deception (Dou et al., 2020), while recent



Graph 1: Evolution of Fraud Detection Methods from Rule-Based Systems to Graph Neural Networks

Table 2: Key Graph Neural Network Techniques Relevant to Healthcare Fraud Detection

Gnn method	*Graph type*	*Fraud detection role*	*Advantage*	*Relevant citation*
GCN	Homogeneous or semi-supervised graph	Learns neighborhood fraud patterns	Strong relational learning	Kipf & Welling (2016)
GraphSAGE	Large inductive graph	Detects new suspicious providers	Scalable to unseen nodes	Hamilton et al. (2017)
GAT	Attention-based graph	Weights important fraud links	Improves interpretability	Veličković et al. (2017)
Heterogeneous GNN	Multi-entity healthcare graph	Models providers, claims, patients, codes	Captures complex claim systems	Hong et al. (2024)
Hierarchical Attention GNN	Attributed healthcare network	Detects layered fraud signals	Learns entity and relation importance	Lu et al. (2023)
Dynamic Graph Fusion	Temporal heterogeneous graph	Tracks evolving fraud behavior	Supports time-aware detection	Zhang et al. (2022)

reviews and applications confirm growing interest in GNN-based medical claims fraud detection (Cheng et al., 2025; Muhammad et al., 2025).

Imbalanced Learning and Cost-Sensitive Fraud Detection

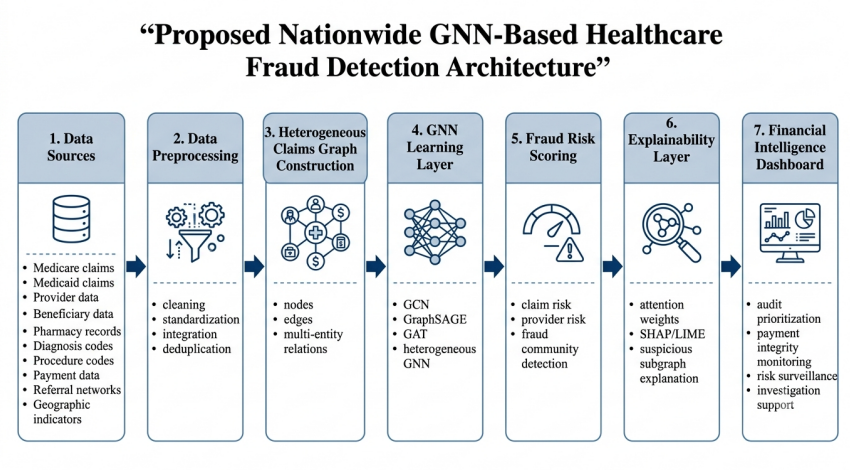
Healthcare fraud datasets are highly imbalanced because confirmed fraud cases are far fewer than legitimate claims. This imbalance can cause models to favor normal claims while missing fraudulent activity. SMOTE can reduce imbalance by creating synthetic minority samples (Chawla et al., 2002), while broader imbalance-learning research highlights the need for robust sampling, weighting, and evaluation strategies (Krawczyk, 2016). In fraud-focused GNNs, imbalance-aware methods improve detection of rare fraud nodes (Liu et al., 2021). Cost-sensitive learning is also important because missing a fraudulent claim may create greater financial harm than incorrectly flagging a legitimate claim (Shi et al., 2023).

Proposed Nationwide GNN-Based Fraud Detection Architecture

System Overview

The proposed architecture is designed as a nationwide healthcare fraud detection and financial risk intelligence system for Medicare and Medicaid programs. Unlike conventional fraud detection systems that examine claims as isolated records, this framework treats public healthcare insurance data as a connected network of providers, beneficiaries, claims, diagnoses, procedures, pharmacies, facilities, locations, and payment flows. This approach is suitable because healthcare fraud often emerges through hidden relationships, repeated service patterns, abnormal billing clusters, coordinated referrals, and suspicious provider-beneficiary interactions rather than through a single claim alone.

The system begins with the integration of multiple data sources, including claims data, provider enrollment



Graph 2: Proposed Nationwide GNN-Based Healthcare Fraud Detection Architecture

files, beneficiary records, procedure codes, diagnosis codes, pharmacy transactions, payment histories, referral records, and geographic indicators. These data are cleaned, standardized, and transformed into a heterogeneous graph in which each healthcare entity is represented as a node and each interaction is represented as an edge. Graph-based anomaly detection is useful in this context because it can identify unusual relational patterns that traditional tabular models may overlook (Akoglu et al., 2015; Chandola et al., 2009). Recent studies also show that graph neural networks are increasingly effective for fraud detection because they learn from both entity attributes and network structure (Dou et al., 2020; Cheng et al., 2025; Hong et al., 2024).

After graph construction, the graph neural network learning layer generates embeddings for providers, beneficiaries, claims, and other entities. These embeddings are then used to classify suspicious claims, rank high-risk providers, detect abnormal referral communities, and support financial risk scoring. The final layer converts model outputs into an explainable financial intelligence dashboard for auditors, regulators, and payment integrity teams.

Healthcare Claims Graph Construction

The healthcare claims graph is structured as a heterogeneous network because Medicare and Medicaid systems contain different types of entities and relationships. Nodes may represent providers, beneficiaries, claims, diagnoses, procedures, pharmacies, facilities, locations, and payment entities. Edges represent interactions such as treatment, billing, referral, prescription, shared diagnosis, shared beneficiary, provider-location association, or payment transfer. This design enables the system to detect fraud signals such as repeated billing, unusually dense provider-beneficiary connections, suspicious pharmacy patterns, abnormal diagnosis-procedure combinations, and geographic outliers.

Feature Engineering and Risk Indicators

The framework uses provider-level, claim-level, beneficiary-level, temporal, geographic, and financial features. Provider-level features include number of claims, average reimbursement, service diversity, referral intensity, and beneficiary concentration. Claim-level features include billed amount, procedure code, diagnosis code, service date, claim frequency, and reimbursement type. Beneficiary-level features include repeated visits, multiple providers, prescription frequency, and unusual service combinations. Temporal features capture billing spikes, repeated claims within short periods, and sudden changes in provider behavior. Geographic features identify unusual travel patterns, cross-state billing, and service clusters. Financial indicators include total reimbursement, high-cost claims, payment concentration, and estimated exposure to improper payments.

GNN Learning Layer

The GNN learning layer transforms the healthcare claims graph into meaningful representations for fraud detection. Graph Convolutional Networks can aggregate information from neighboring nodes to identify suspicious local patterns (Kipf & Welling, 2016). GraphSAGE supports inductive learning, allowing the model to generalize to new providers, beneficiaries, and claims that were not present during training (Hamilton et al., 2017). Graph Attention Networks assign different weights to neighboring nodes, helping the model focus on the most suspicious relationships (Veličković et al., 2017). Heterogeneous GNNs are especially suitable because Medicare and Medicaid claims involve multiple entity types and relation types (Lu et al., 2023; Zhang et al., 2022). Through these methods, the system can detect suspicious claims, risky providers, hidden fraud communities, and financial exposure across nationwide healthcare insurance networks.

Table 3: Proposed Heterogeneous Medicare-Medicaid Claims Graph Schema

Node type	*Edge type*	*Example relationship*	*Fraud signal*	*Analytical use*
Provider, Beneficiary	Treatment edge	Provider treats beneficiary	Repeated visits or excessive services	Detect abnormal provider-patient patterns
Provider, Claim	Billing edge	Provider submits claim	High claim frequency or unusual billing volume	Provider-level fraud scoring
Claim, Diagnosis	Diagnosis edge	Claim linked to diagnosis code	Diagnosis not matching procedure	Detect coding abuse or upcoding
Provider, Provider	Referral edge	Provider refers to another provider	Closed referral loops	Detect collusive provider networks
Pharmacy, Beneficiary	Prescription edge	Pharmacy serves beneficiary	Repeated high-cost prescriptions	Identify prescription fraud risk
Provider, Location	Geographic edge	Provider operates in a location	Unusual distance or service concentration	Detect geographic anomalies

Methodological Design and Model Development

Research Design

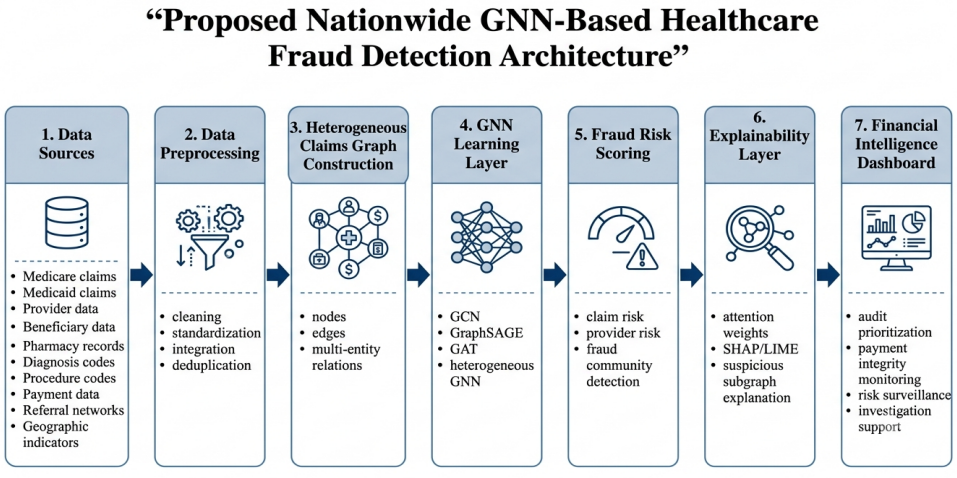
This paper adopts a conceptual and technical architecture research design to propose a scalable graph neural network-based framework for nationwide healthcare fraud detection and financial risk intelligence in Medicare and Medicaid systems. Rather than presenting a single experimental dataset, the study develops a structured model that integrates prior evidence from healthcare fraud detection, graph-based anomaly detection, imbalanced learning, explainable artificial intelligence, and financial risk analytics. This approach is suitable because Medicare and Medicaid fraud is not only a claim-level problem but also a relational and system-level problem involving providers, beneficiaries, diagnoses, procedures, prescriptions, payment flows, and referral networks. Existing studies have shown that conventional fraud detection models are useful but often limited when fraud behavior is hidden within complex relationships and coordinated networks (Bolton & Hand, 2002; Chandola et al., 2009; Akoglu et al., 2015). Therefore, the proposed framework positions graph neural networks as a national intelligence architecture capable of learning from both healthcare attributes and network structures.

Data Sources and Integration Strategy

A real-world deployment of the proposed system would require integration of multiple administrative, clinical, financial, and investigative data sources. These may include Medicare claims, Medicaid claims, provider enrollment records, beneficiary demographic records, diagnosis codes, procedure codes, pharmacy and prescription records, payment histories, facility information, referral relationships, geographic indicators, and confirmed fraud investigation outcomes. These datasets would be transformed into a unified heterogeneous healthcare claims graph where different entities are represented as nodes and their relationships are represented as edges. For example, a provider may be connected to claims, procedures, beneficiaries, pharmacies, and locations, while beneficiaries may be connected to diagnoses, prescriptions, and service utilization patterns. This integrated structure allows the model to detect suspicious patterns such as abnormal billing frequency, repeated provider-beneficiary interactions, unusual diagnosis-procedure combinations, excessive referrals, and geographically inconsistent service claims. Before model training, privacy, compliance, and governance controls must be implemented to protect sensitive health information. These controls should include de-identification, role-based access, data minimization, audit logging, secure storage, and compliance with healthcare data protection requirements.

Model Training and Fraud Classification

The proposed model supports multiple fraud classification tasks. First, binary classification can be used to distinguish fraudulent or high-risk claims from legitimate claims. Second, multi-class classification can be used to categorize fraud types, such as upcoding, phantom billing, unnecessary services, prescription abuse, provider collusion, or abnormal referral behavior. Third, provider-level risk classification can assign risk scores to providers based on billing behavior, network centrality, claim patterns, and relationships with other high-risk entities. Fourth, claim-level anomaly detection can identify individual claims that deviate from expected clinical, financial, or geographic patterns. Fifth, network-level detection can identify suspicious communities, fraud rings, or coordinated provider-beneficiary clusters. Where confirmed fraud labels are available, supervised GNN training can be applied. However, because confirmed fraud labels are often incomplete or delayed, semi-supervised learning is also necessary. Semi-supervised GNNs can learn from a



Graph 3: Model Workflow for Fraud Risk Scoring and Explainable Audit Review

small number of labeled fraud cases while using the larger unlabeled claims graph to detect hidden risk patterns (Kipf & Welling, 2016; Hamilton et al., 2017; Wu et al., 2020).

Handling Class Imbalance

Healthcare fraud detection is naturally affected by severe class imbalance because fraudulent claims represent a very small proportion of total claims. If this issue is ignored, the model may achieve high overall accuracy while failing to detect rare but financially damaging fraud cases. To address this, the proposed architecture should combine oversampling, cost-sensitive loss functions, focal learning, anomaly scoring, and imbalance-aware GNN sampling. SMOTE can be used to synthetically increase minority fraud examples during training (Chawla et al., 2002). Cost-sensitive learning can assign higher penalties to missed fraud cases, making the model more sensitive to high-risk events (Shi et al., 2023). Imbalance-aware GNN strategies can also improve fraud detection by selecting informative neighbors, reducing majority-class dominance, and improving learning from rare fraudulent structures (Krawczyk, 2016; Liu et al., 2021).

Explainability and Audit Support

Explainability is essential because healthcare fraud detection affects providers, beneficiaries, auditors, and public funds. A fraud risk score should not function as a hidden decision but as an evidence-supported signal for human review. The proposed system should therefore include SHAP values, LIME explanations, graph attention weights, suspicious subgraph extraction, and risk factor ranking. SHAP can identify which variables contributed most to a prediction, while LIME can provide local explanations for individual claims or providers (Ribeiro et al., 2016; Lundberg & Lee, 2017). Attention weights can show which neighboring nodes or relationships influenced the GNN output, while suspicious subgraph extraction can reveal connected providers, beneficiaries, claims, procedures, or pharmacies involved in abnormal activity. This strengthens audit transparency and supports reliable forensic interpretation, which is important because fraud evidence must be valid, explainable, and reviewable (Stern et al., 2019; Hancock et al., 2023).

Expected Results, Evaluation Metrics, and Financial Risk Intelligence

Expected Fraud Detection Outcomes

The proposed Graph Neural Network-based nationwide healthcare fraud detection system is expected to produce stronger fraud intelligence outcomes than conventional claim-level fraud detection models because it analyzes both healthcare attributes and relational structures. Traditional fraud detection methods often evaluate claims independently, using billing amount, diagnosis code, procedure code, provider specialty, or beneficiary profile as separate variables. While these methods can identify obvious irregularities, they may fail to detect coordinated and hidden fraud patterns that emerge across provider networks, referral chains, shared beneficiaries, pharmacy relationships, and geographic billing clusters. Since healthcare fraud is frequently relational, adaptive, and network-based, the proposed GNN model is expected to identify suspicious providers, abnormal billing communities, high-risk claims, coordinated fraud rings, and unusual payment flows more effectively than isolated models.

At the provider level, the system is expected to detect unusual billing behaviors such as excessive claim volume, repetitive high-cost procedures, abnormal diagnosis-procedure combinations, unusually high reimbursement intensity, or billing patterns that differ sharply from comparable providers. At the claim level, the model can assign risk scores to claims that appear inconsistent with patient history, provider behavior, diagnosis patterns, pharmacy activity, or geographic service expectations. At the network level, the GNN can identify clusters of providers, beneficiaries, pharmacies, or facilities that repeatedly interact in ways that suggest collusion, referral manipulation, phantom billing, or systematic upcoding. This reflects the value of graph-based anomaly detection, where suspicious behavior is not only defined by individual records but also by unusual relationships and structural patterns within the network (Akoglu et al., 2015; Chandola et al., 2009).

The proposed system is also expected to improve detection of camouflaged fraudsters who attempt to imitate normal behavior. GNNs are suitable for this task because they learn from node attributes and neighborhood information, allowing the model to compare a provider's behavior with the behavior of connected entities and peer groups. Prior GNN fraud studies show that relational learning can strengthen fraud detection where malicious actors hide within legitimate transaction networks (Dou et al., 2020; Cheng et al., 2025). In Medicare and Medicaid systems, this capability is important because fraudulent activity may be distributed across multiple entities rather than concentrated in one claim.

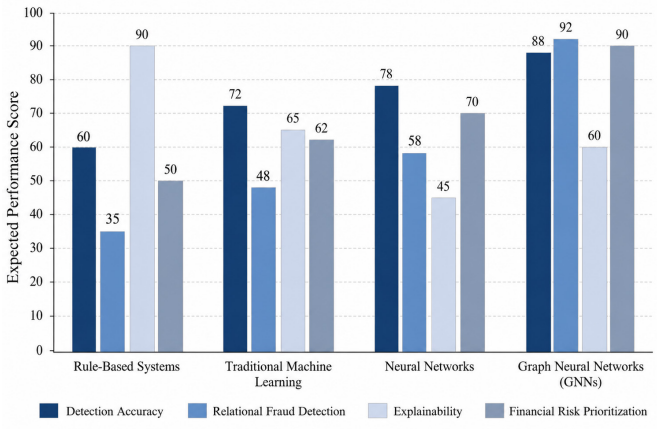
Evaluation Metrics

The performance of the proposed model should be evaluated using both statistical detection metrics and financial intelligence metrics. Precision measures the proportion of flagged cases that are truly fraudulent or suspicious. High precision is important because excessive false alerts can waste investigative resources and unfairly burden legitimate providers. Recall measures the proportion of actual fraud cases detected by the model. High recall is necessary because missed fraud can lead to continued financial loss. The F1-score balances precision and recall, making it useful when healthcare fraud datasets are highly imbalanced, as confirmed fraud cases are usually much fewer than legitimate claims (Chawla et al., 2002; Krawczyk, 2016).

AUROC measures the model's ability to separate fraudulent and legitimate cases across classification

Table 4: Evaluation Metrics for Nationwide Healthcare Fraud Detection and Financial Risk Intelligence

Metric	*Meaning*	*Importance for fraud detection*	*Financial intelligence value*
Precision	Share of flagged cases that are truly suspicious	Reduces unnecessary investigations	Improves audit resource allocation
Recall	Share of actual fraud cases detected	Minimizes missed fraud	Prevents continued financial leakage
F1-score	Balance between precision and recall	Useful for imbalanced fraud datasets	Supports balanced enforcement decisions
AUROC	Separates fraud and non-fraud cases	Measures general classification strength	Supports model comparison
AUPRC	Precision-recall performance for rare fraud	Strong metric for imbalanced data	Identifies high-value suspicious cases
False positive rate	Legitimate cases incorrectly flagged	Protects compliant providers	Reduces wasted audit cost
Detection latency	Speed of fraud identification	Enables early intervention	Supports payment integrity
Cost savings	Estimated prevented or recovered losses	Shows practical fraud reduction	Measures financial return
Investigation yield	Useful audit outcomes from flagged cases	Tests operational effectiveness	Improves enforcement productivity
Explainability usefulness	Clarity of model explanations	Supports auditor trust	Strengthens evidence-based review



Graph 4: Comparative Expected Performance of Fraud Detection Models

thresholds, while AUPRC is especially useful for imbalanced fraud detection because it focuses on precision-recall performance in rare-event settings. False positive rate should be monitored to ensure that legitimate providers are not repeatedly flagged without sufficient evidence. Detection latency measures how quickly the system can identify suspicious behavior after a claim is submitted or paid. Cost savings estimate the financial value of prevented, recovered, or prioritized fraudulent payments. Investigation yield measures the proportion of flagged cases that result in meaningful audit findings, payment holds, recoveries, or enforcement actions. Explainability usefulness evaluates whether auditors can understand why a claim, provider, or network was flagged. This is essential because fraud detection systems must support evidence-based review rather than operate as opaque decision tools (Ribeiro et al., 2016; Lundberg & Lee, 2017; Hancock et al., 2023).

Financial Risk Intelligence Layer

The financial risk intelligence layer transforms model outputs into practical decision support for Medicare and Medicaid oversight. Instead of simply classifying claims as fraudulent or legitimate, the system generates risk scores, provider risk profiles, suspicious subgraph summaries, payment exposure estimates, and audit prioritization lists. This allows public insurance agencies to rank claims and providers according to financial risk, urgency, network suspiciousness, and expected recovery value.

The system can support payment integrity by flagging suspicious claims before payment, identifying providers requiring closer monitoring, and detecting abnormal payment flows across regions. It can also support claim prioritization by directing auditors toward cases with high fraud probability and high financial exposure. For investigation triage, the model can group related claims, providers, beneficiaries, and pharmacies into explainable fraud-risk networks. This helps investigators understand whether a suspicious claim is isolated or part of a broader pattern. The system can also enable early warning surveillance by monitoring emerging fraud clusters across states, provider types, and service categories.

Comparative Performance Expectation

Compared with rule-based systems, traditional machine learning, and standard neural networks, GNNs are expected to provide stronger performance in detecting relational healthcare fraud. Rule-based systems are transparent but rigid, making them weak against adaptive fraud. Traditional machine learning can detect statistical anomalies but often depends on manually engineered features. Neural networks can model complex nonlinear patterns but may still treat claims as isolated records. In contrast, GNNs model both attributes and relationships, making them more effective for detecting hidden, coordinated, and community-based fraud patterns (Kipf & Welling, 2016; Hamilton et al., 2017; Veličković et al., 2017; Wu et al., 2020).

Discussion, Policy Implications, and Ethical Considerations

Discussion of Technical Significance

The proposed graph neural network-based framework offers a technically advanced approach to healthcare fraud detection because it moves beyond isolated claim-level analysis and captures the relational nature of fraud. Traditional fraud detection models often rely on tabular features such as claim amount, billing frequency, provider specialty, or procedure type. While useful, these methods may fail to identify hidden connections among providers, beneficiaries, pharmacies, facilities, diagnosis codes, procedure codes, and payment patterns. GNNs are more suitable because they learn from both node attributes and network relationships, allowing the model to detect suspicious structures such as repeated provider-beneficiary interactions, abnormal referral communities, shared billing patterns, and coordinated fraud rings. This is particularly important for Medicare and Medicaid systems, where claims are generated across large, multi-entity and multi-state networks. By learning structural patterns from heterogeneous healthcare graphs, the proposed system can support provider-level risk prediction, claim-level anomaly detection, and network-level fraud intelligence.

Policy and Governance Implications

From a policy perspective, a nationwide GNN-based fraud detection system can strengthen public insurance program integrity by enabling earlier detection of fraud, waste, and abuse. Federal and state agencies responsible for Medicare and Medicaid oversight can use the system to prioritize audits, monitor high-risk provider networks, and allocate investigative resources more efficiently. Instead of treating all claims equally, the model can generate risk scores that help agencies focus on claims or providers with the strongest indicators of financial irregularity. This can improve payment integrity, reduce improper payments, and support more evidence-based enforcement decisions. The framework can also help policymakers understand emerging fraud patterns across regions, specialties, and provider networks, making it useful for long-term healthcare financial risk intelligence.

Ethical, Legal, and Operational Concerns

Despite its potential, the proposed system must be implemented with strong ethical and legal safeguards. A major concern is the risk of false positives, where legitimate providers may be incorrectly flagged as suspicious. Such errors can damage reputations, delay reimbursements, and create unnecessary administrative burdens. Therefore, AI-generated risk scores should not be used as automatic proof of fraud. Instead, they should serve as decision-support tools for trained investigators. Provider fairness, model transparency, explainability, and audit accountability are also essential. Explainable AI methods should be used to show why a claim, provider, or network was flagged. Privacy protection is equally important because Medicare and Medicaid data contain sensitive health and financial information. Strong governance, secure data handling, and compliance with healthcare privacy regulations are necessary to maintain trust.

Deployment Challenges

Several deployment challenges may affect nationwide implementation. Medicare and Medicaid datasets differ across states, systems, and administrative structures, creating interoperability problems. Medicaid variation across states may also affect model consistency and generalizability. Other challenges include limited confirmed fraud labels, evolving fraud strategies, computational scalability, privacy restrictions, and institutional readiness. Fraudsters may adapt once detection systems become stronger, requiring continuous model updating and monitoring. For successful deployment, the framework should combine scalable infrastructure, cross-agency governance, human oversight, fairness auditing, and ongoing validation before being used in real-world enforcement settings.

Conclusion

This paper concludes that graph neural networks provide a strong and scalable pathway for nationwide healthcare fraud detection and financial risk intelligence in Medicare and Medicaid systems. Healthcare fraud is rarely an isolated event because fraudulent behavior often emerges through complex relationships among providers, beneficiaries, claims, pharmacies, diagnoses, procedures, referral patterns, payment flows, and geographic locations. Traditional rule-based systems and conventional machine learning models are useful for detecting known fraud patterns, but they often struggle to identify hidden, adaptive, and network-based fraud schemes. In contrast, graph neural networks are highly suitable for this problem because they can learn from both entity attributes and relational structures within large healthcare claims networks.

The proposed framework demonstrates how Medicare and Medicaid data can be transformed into a heterogeneous claims graph that supports provider-level, claim-level, and network-level fraud detection. By connecting multiple data sources into a graph intelligence layer, the system can detect suspicious billing behaviors, unusual referral clusters, abnormal provider-beneficiary relationships, high-risk procedure combinations, and coordinated fraud communities. This aligns with prior studies on graph-based anomaly detection, healthcare fraud analytics, and GNN-based fraud detection, which emphasize the importance of relational learning in complex fraud environments (Akoglu et al., 2015; Dou et al., 2020; Hong et al., 2024; Muhammad et al., 2025). Therefore, the study reaffirms that GNNs can reveal hidden connections that may be missed by traditional fraud detection models.

Future Research

Future research should focus on real-world validation using large-scale Medicare and Medicaid datasets. Since healthcare fraud patterns vary across states, provider types, clinical specialties, and reimbursement systems, future studies should test the proposed architecture across diverse public insurance environments. Dynamic graph learning should also be explored to capture changing fraud behaviors over time, including sudden billing spikes, emerging provider networks, and evolving referral relationships. In addition, federated GNN training should be investigated as a privacy-preserving approach that allows multiple healthcare agencies or state Medicaid programs to collaborate without transferring raw patient or claims data.

Further research should also examine privacy-preserving analytics, secure aggregation, differential privacy, and compliance-aware model deployment. Fairness auditing is another important direction because fraud detection systems must avoid unfairly targeting specific provider groups, geographic regions, or vulnerable beneficiary populations. Explainable fraud dashboards should be developed to help investigators understand why a provider, claim, or network was flagged as suspicious. These dashboards should include risk scores, influential features, suspicious subgraphs, payment exposure, and audit-priority recommendations. Finally, future work should examine how GNN-based fraud intelligence can be integrated into government investigation workflows, payment integrity units, and healthcare oversight systems.

Final Contribution Statement

The major contribution of this paper is the development of a scalable, explainable, and policy-relevant GNN architecture for nationwide healthcare fraud detection and financial risk intelligence. The framework connects healthcare claims analytics, fraud detection, financial risk scoring, explainable AI, and public insurance governance into one unified model. By combining relational learning with financial intelligence, the proposed system offers a stronger foundation for protecting Medicare and Medicaid resources, improving audit prioritization, and supporting more transparent healthcare program oversight.

References

[1] Akoglu, L., Tong, H., & Koutra, D. (2015). Graph based anomaly detection and description: a survey. Data mining and knowledge discovery, 29(3), 626-688.

[2] Bauder, R. A., & Khoshgoftaar, T. M. (2017, December). Medicare fraud detection using machine learning methods. In 2017 16th IEEE international conference on machine learning and applications (ICMLA) (pp. 858-865). IEEE.

[3] Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. Statistical science, 17(3), 235-255.

[4] Kalange, O. S., Kahat, R. S., Kale, A. S., Kale, T. R., & Joglekar, P. S. (2022). Implementation of Various Machine Learning Algorithms for Traffic Sign Detection and Recognition.

[5] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM computing surveys (CSUR), 41(3), 1-58.

[6] 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.

[7] Cheng, D., Zou, Y., Xiang, S., & Jiang, C. (2025). Graph neural networks for financial fraud detection: a review. Frontiers of Computer Science, 19(9), 199609.

[8] Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., & Yu, P. S. (2020, October). Enhancing graph neural network-based fraud detectors against camouflaged fraudsters. In Proceedings of the 29th ACM international conference on information & knowledge management (pp. 315-324).

[9] du Preez, A., Bhattacharya, S., Beling, P., & Bowen, E. (2025). Fraud detection in healthcare claims using machine learning: A systematic review. Artificial Intelligence in Medicine, 160, 103061.

[10] Stern, H. S., Cuellar, M., & Kaye, D. (2019). Reliability and validity of forensic science evidence. Significance, 16(2), 21-24.

[11] Hancock, J. T., Bauder, R. A., Wang, H., & Khoshgoftaar, T. M. (2023). Explainable machine learning models for Medicare fraud detection. Journal of Big Data, 10(1), 154.

[12] Hamilton, W., Ying, Z., & Leskovec, J. (2017). Inductive representation learning on large graphs. Advances in neural information processing systems, 30.

[13] Hong, B., Lu, P., Xu, H., Lu, J., Lin, K., & Yang, F. (2024). Health insurance fraud detection based on multi-channel heterogeneous graph structure learning. Heliyon, 10(9).

[14] Johnson, J. M., & Khoshgoftaar, T. M. (2019). Medicare fraud detection using neural networks. Journal of Big Data, 6(1), 63.

[15] Joudaki, H., Rashidian, A., Minaei-Bidgoli, B., Mahmoodi, M., Geraili, B., Nasiri, M., & Arab, M. (2014). Using data mining to detect health care fraud and abuse: a review of literature. Global journal of health science, 7(1), 194.

[16] Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.

[17] Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. Progress in artificial intelligence, 5(4), 221-232.

[18] Liu, Y., Ao, X., Qin, Z., Chi, J., Feng, J., Yang, H., & He, Q. (2021, April). Pick and choose: a GNN-based imbalanced learning approach for fraud detection. In Proceedings of the web

conference 2021 (pp. 3168-3177).
[19] Lu, J., Lin, K., Chen, R., Lin, M., Chen, X., & Lu, P. (2023). Health insurance fraud detection by using an attributed heterogeneous information network with a hierarchical attention mechanism. BMC Medical Informatics and Decision Making, 23(1), 62.
[20] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.
[21] Muhammad, R., Tbaishat, D., Nazir, A., Yacoub, S., AbdulRazek, M., El-Enen, M. A. A., & Sahlol, A. T. (2025). Fraud detection and explanation in medical claims using GNN architectures. Scientific Reports.
[22] Naidoo, K., & Marivate, V. (2020, April). Unsupervised anomaly detection of healthcare providers using generative adversarial networks. In Conference on e-Business, e-Services and e-Society (pp. 419-430). Cham: Springer International Publishing.
[23] Ortega, P. A., Figueroa, C. J., & Ruz, G. A. (2006). A Medical Claim Fraud/Abuse Detection System based on Data Mining: A Case Study in Chile. DMIN, 6, 26-29.
[24] Pereira, R. D., & Murai, F. (2021). How effective are Graph Neural Networks in Fraud Detection for Network Data?. arXiv preprint arXiv:2105.14568.
[25] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).
[26] Shi, H., Tayebi, M. A., Pei, J., & Cao, J. (2023). Cost-sensitive learning for medical insurance fraud detection with temporal information. IEEE Transactions on Knowledge and Data Engineering, 35(10), 10451-10463.
[27] Thornton, D., Brinkhuis, M., Amrit, C., & Aly, R. (2015). Categorizing and describing the types of fraud in healthcare. Procedia Computer Science, 64, 713-720.
[28] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. arXiv preprint arXiv:1710.10903.
[29] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020). A comprehensive survey on graph neural networks. IEEE transactions on neural networks and learning systems, 32(1), 4-24.
[30] Zhang, J., Yang, F., Lin, K., & Lai, Y. (2022). Hierarchical multi-modal fusion on dynamic heterogeneous graph for health insurance fraud detection. In 2022 IEEE International Conference on Multimedia and Expo (ICME) (pp. 1–6). IEEE. https://doi.org/10.1109/ICME52920.2022.9859871