International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(20s) 2026 pp. 224-234
https://doi.org/10.70917/ijcisim-2026-5176
Article

A Hybrid, Privacy-Preserving AI Framework for 340B Program Fraud Detection: Moving Beyond Rule-Based Audits with Federated Learning and Graph Neural Networks

Pinaki Bose

Independent Researcher, India
pinaki.investing@gmail.com

Abstract: The 340B Drug Pricing Program, critical for supporting safety-net healthcare providers, faces significant challenges from fraud, waste, and abuse, particularly through drug diversion and duplicate discounts. Current program integrity relies heavily on static, rule-based audits, which are increasingly ineffective against novel and sophisticated, collusive fraud schemes. This paper proposes a novel, multi-layered artificial intelligence (AI) framework designed to move beyond this reactive paradigm. The framework combines: (1) unsupervised transactional anomaly detection, utilizing Autoencoders and Isolation Forests to identify entity-level outliers; and (2) relational fraud detection, using an inductive Graph Neural Network (GNN), GraphSAGE, to model the complex relationships between covered entities, contract pharmacies, and providers to uncover collusive networks. Crucially, this hybrid model is designed within a privacy-preserving architecture. We propose the use of Federated Learning (FL) and Differential Privacy (DP) to train the anomaly detection modules collaboratively, allowing participating entities to build robust, global models without centralizing or exposing sensitive HIPAA-protected patient data. The framework integrates Explainable AI (XAI) methodologies (SHAP and GNNExplainer) for regulatory auditability and a concept drift detection component with a Human-in-the-Loop (HITL) active learning loop to ensure long-term adaptability to evolving fraud tactics. This approach provides a blueprint for a proactive, adaptive, and privacy-compliant system to safeguard 340B program integrity.
Keywords: 340B Drug Pricing Program, Fraud Detection, Machine Learning, Graph Neural Networks (GNN), Anomaly Detection, Federated Learning, Privacy-Preserving AI, Rule-Based Systems, Healthcare Compliance, Explainable AI (XAI)

1. Introduction

1.1 The 340B Drug Pricing Program: Purpose and Scale

Enacted in 1992 as part of the Veterans Health Care Act, the 340B Drug Pricing Program is a cornerstone of the U.S. healthcare safety net [1]. The program mandates that pharmaceutical manufacturers participating in Medicaid provide outpatient drugs to eligible "covered entities" (CEs) at deeply discounted prices [1]. These CEs, which include Disproportionate Share Hospitals (DSH), community health centers, and other safety-net providers, can then generate savings, which congressional language intended to "stretch scarce federal resources as far as possible, reaching more eligible patients and providing more comprehensive services".

The program's scale is immense. By 2025, 340B purchases comprise over 13% of the total U.S. prescription drug market. Participating hospitals and CEs can achieve average savings of 25% to 50% on outpatient drugs. This financial scale, coupled with the program's complexity, creates significant financial incentives [2] and, consequently,

a vast attack surface for fraud, waste, and abuse [3]. Losses due to healthcare fraud are estimated to be in the tens of billions of dollars annually, with some estimates as high as 10% of total healthcare expenditures.

1.2 The 340B Fraud Landscape: Diversion and Duplicate Discounts

Program integrity hinges on adherence to two primary statutory prohibitions. The first is drug diversion, defined as the resale or transfer of 340B-discounted drugs to individuals who are not eligible patients of the covered entity [1]. The second is the prohibition against duplicate discounts, which occurs when a manufacturer provides an upfront 340B discount to a CE for a drug that is also subject to a back-end Medicaid rebate claim from a state agency [1].

The 340B ecosystem is a complex web of stakeholders with competing incentives, including CEs, drug manufacturers, a rapidly expanding network of contract pharmacies (CPs), state Medicaid agencies, and the program administrator, the Health Resources and Services Administration (HRSA) [1]. This complexity makes detecting diversion and duplicate discounts exceptionally difficult.

1.3 The Inadequacy of Static Rule-Based Audits

Current 340B program integrity is primarily enforced through audits conducted by HRSA and, separately, by manufacturers, who are permitted to audit CEs to ensure compliance. These audits rely on traditional, rule-based systems [3]. This paradigm is fundamentally ill-equipped to manage the modern fraud landscape for several reasons:

1. **Rigidity and Reactivity:** Rule-based systems depend on explicit, predefined "if-then" logic. They are inherently reactive and rigid, incapable of identifying novel or evolving fraud patterns [4]. As fraudsters, who now employ sophisticated tactics like synthetic identities and even AI-driven attacks, learn the rules, they adapt their methods to circumvent them.

2. **Manual Intensity and False Positives:** These systems are resource-intensive, often requiring slow, manual audits. Their rigidity also leads to a high volume of false positives, flagging legitimate transactions that require costly manual review.

3. **Focus on Administrative Compliance:** HRSA audit findings often focus on straightforward administrative errors, such as an "Incorrect accurate 340B OPAIS Record," rather than complex, behavioral fraud.

This systemic weakness creates a "compliance paradox." A 2025 report from the American Hospital Association (AHA) highlights that 340B hospital audits show declining findings of diversion and duplicate discounts, suggesting very high rates of compliance. This, however, stands in stark contrast to the persistent, multi-billion-dollar estimates of healthcare fraud and academic studies indicating "margin-motivated behavior" by participating entities.

These findings are not contradictory. The high compliance reported is likely an artifact of an inadequate audit process. CEs are "teaching to the test", that is, they are successfully complying with the static, predictable rules being audited. The legacy rule-based systems [4] are simply failing to detect sophisticated, collusive, or novel fraud, which is where the most significant financial losses occur. This detection gap is the primary justification for a new technological paradigm.

1.4 Paper Contribution: A Hybrid, Privacy-Preserving Framework

This paper proposes a novel, multi-layered AI framework to address the limitations of rule-based systems. The framework is designed to detect both anomalous transactional behavior at the entity level and sophisticated, collusive fraud at the network level. Its core contributions are:

1. **A Hybrid AI Model:** A two-module system combining unsupervised anomaly detection (e.g., Autoencoders [5], Isolation Forests [6]) with relational Graph Neural Networks (GNNs) [7].

2. **A Privacy-Preserving Architecture:** An architecture integrating Federated Learning (FL) and Differential Privacy (DP) to enable collaborative model training without centralizing or exposing sensitive patient data, thereby adhering to HIPAA and HITECH Act requirements [8].

3. **An Interpretable and Adaptive System:** A design incorporating Explainable AI (XAI) [9, 10] for auditability and concept drift detection for long-term robustness against adversarial attacks.

Table I provides a comparative analysis of the traditional and proposed methodologies.

TABLE I. Comparison of 340B Audit Methodologies

Feature	**Traditional Rule-Based Audits (TRBA)**	**Proposed AI Framework (PAF)**
Detection Method	Static IF-THEN logic, data matching	Behavioral and relational pattern analysis
Novel Fraud Detection	Incapable; new rules must be manually added [4]	Designed to identify novel, unknown patterns
Data Handling	Manual review, siloed datasets	Federated, holistic network analysis
Key Weakness	High false positives, rigidity, adversarial evasion	Model and data complexity, computational cost
Primary Focus	Record-keeping compliance (e.g., OPAIS accuracy)	Collusive and behavioral fraud (e.g., diversion rings)

2. Data Foundations and Enrichment Strategies

2.1 The Public Data Ecosystem: HRSA OPAIS

The foundational dataset for any 340B analysis is the public HRSA 340B Office of Pharmacy Affairs Information System (OPAIS) [11], [12]. This system provides publicly downloadable files, notably the "Covered Entity Daily Report," available in both Microsoft Excel and JSON formats [13], [14].

This report is not a single flat file but a relational structure composed of three primary worksheets (or JSON objects):

1. **'Covered Entities' Worksheet:** Contains entity-level data, including the 340B ID (the unique identifier), CE ID, Entity Type (e.g., DSH, CHC, RRC), Participating (a true/false active status indicator), and Participating Start Date [14].

2. **'Contract Pharmacies' Worksheet:** Defines the relationship between CEs and their CPs, following the registration structure documented in [15]. Key fields include the 340B ID (linking to the CE), Pharmacy Name, Pharmacy Address 1, Pharmacy Address 2, Pharmacy City, Pharmacy State, Pharmacy Zip, and Contract Termination Date [14].

3. **'Shipping Addresses' Worksheet:** Lists all approved shipping locations. Fields include the 340B ID (linking to the CE), Organization, Address Line 1, City, State, Zip, and a critical flag, 340B ID Street Address, which indicates if the address is the primary CE location or an auxiliary shipping site [14].

Critically, the OPAIS data contains no transactional, patient, or financial information. Its value is as a graph scaffold: it explicitly defines the nodes (CEs, CPs, Shipping Locations) and their declared "contract" and "ships_to" relationships.

2.2 Essential Private Data (The Missing Link)

To detect behavioral fraud, the public OPAIS scaffold must be integrated with the private, sensitive, and high-volume transactional data firewalled within each CE and its partners [3]. The framework proposed in this paper assumes access, via the privacy-preserving methods detailed in Section 4, to the following data categories:

1. **Dispensing / Prescription Data:** (e.g., Hashed Patient_ID, Prescriber_ID, Drug_NDC, Quantity, Date_Filled, Pharmacy_ID).

2. **Claims Data:** (e.g., Claim_ID, Hashed Patient_ID, Provider_ID, Diagnosis_Code (ICD-10), Procedure_Code (CPT), Amount_Billed, Service_Date) [3].

3. **Electronic Health Record (EHR) Data:** (e.g., Hashed Patient_ID, Encounter_Date, Patient_Status (e.g., outpatient, inpatient)) necessary to verify patient eligibility against HRSA's "Patient Definition Guidelines".

2.3 Novel Feature Engineering: Entity Plausibility Scoring

A key innovation of this framework is a data enrichment step that assesses the plausibility of the OPAIS entities before modeling.

1. **Geospatial Validation:** The Pharmacy Address and Shipping Address fields from OPAIS will be cross-referenced against external geospatial and address validation APIs (e.g., the USPS Address Matching System (AMS) API, Google Address Validation API, or Esri Geocoding services). A rule-based check merely validates if an address exists. This AI-driven approach assesses plausibility. These APIs [16] can return metadata such as whether an address is "Commercial" or "Residential," or if it is a known mail-forwarding service. This allows for the creation of a "Geospatial Plausibility Score." A contract pharmacy registered in OPAIS that resolves to a residential address or a P.O. box is a significant red flag for a potential shell company facilitating diversion.

2. **Entity Resolution:** The framework will link OPAIS entity names (CE Name, Pharmacy Name) with external data sources, such as state-level business registries or National Provider Identifier (NPI) databases. This enrichment creates features like Business_Status (e.g., Active, Dissolved) and License_Status (e.g., Valid, Revoked), providing further signals for entity risk.

The complete data requirements are summarized in Table II.

TABLE II. Data Requirements and Enrichment Framework

Data Source	Data Type	Key Fields	Analytical Purpose	Privacy Constraint
HRSA OPAIS CE File	Public, Relational	340B ID, Entity Type, Participating Start Date	Graph Node Creation (CE)	Public
HRSA OPAIS CP File	Public, Relational	340B ID, Pharmacy Name, Pharmacy Address, Contract Termination Date	Graph Node Creation (CP), Graph Edge Creation	Public
Private Claims Data	Private, Transactional	Patient_ID (hashed), Provider_ID, Diagnosis_Code, Claim_Amount	Anomaly Feature Engineering	HIPAA / HITECH
Private Rx Data	Private, Transactional	Patient_ID (hashed), Prescriber_ID, Drug_NDC, Quantity, Pharmacy_ID	Anomaly Feature Engineering, Graph Edge Creation	HIPAA / HITECH
Geospatial APIs	Public, Enrichment	Address_Type (Comm./Res.), Is_PO_Box	Entity Plausibility Scoring	Public
State / NPI Registries	Public, Enrichment	Business_Status (Active/Dissolved), License_Status	Entity Plausibility Scoring	Public

3. Hybrid AI Framework for Fraud Detection

The proposed framework consists of two primary modules designed to operate at different scopes: the entity level and the network level.

3.1 Module 1: Unsupervised Transactional Anomaly Detection (Entity-Level)

Rationale: This module addresses fraud perpetrated by individual anomalous entities (e.g., a single provider engaging in phantom billing, a non-compliant CE). The approach must be unsupervised due to the severe class imbalance (fraud is rare) and the notorious scarcity of reliable, labeled "fraud" data in real-world healthcare datasets [17].

Method 1: Autoencoders (AE) for Pattern Anomaly. A deep learning Autoencoder (AE) or Variational Autoencoder (VAE) [5] is proposed.

- **Training:** The AE is trained on a large corpus of (presumed) legitimate CE claims and prescription data [18]. The model learns a compressed, latent representation of "normal" billing, prescribing, and dispensing behavior.

- **Inference:** When a new transaction or a provider's aggregated profile is fed into the model, its reconstruction error is calculated. A high reconstruction error signifies that the model "struggles" to reconstruct this data, as it deviates significantly from the "normal" patterns it learned [18].

- **Use Case:** This is effective at detecting complex anomalies like "phantom billing", billing for unprovided services, or unusual profiles, such as a prescriber's sudden shift to anomalous diagnosis-procedure code combinations.

Method 2: Isolation Forest (iForest) for Outlier Detection. As a complementary and computationally efficient method, an Isolation Forest (iForest) is included [6].

- **Mechanism:** iForest is an ensemble model that isolates anomalies by building random decision trees. Anomalous data points are "easier" to isolate (require fewer splits) and thus have a shorter average path length across the ensemble of trees.

- **Use Case:** iForest is highly effective and efficient in high-dimensional datasets [6]. It is ideal for a rapid, initial screening of millions of claims to identify extreme outliers (e.g., a prescription with an impossible Quantity, a claim with an extreme Amount_Billed) that a deep AE might "over-normalize."

The primary output of Module 1 is a new, engineered feature: an Anomaly_Score for each entity (provider, pharmacy) and/or transaction. This score can be used to trigger immediate audits or, more powerfully, as a node feature in Module 2.

3.2 Module 2: Relational Fraud Detection via Graph Neural Networks (Network-Level)

Rationale: This module targets sophisticated, collusive fraud, which is the key weakness of traditional audits. Rule-based systems and standard machine learning models fail by treating transactions as independent events. 340B diversion, however, is fundamentally a network problem involving coordinated, non-obvious behavior between CEs, providers, and pharmacies, a structure for which graph-based detection has been applied to healthcare claims specifically [19]. Graph Neural Networks (GNNs) are uniquely capable of modeling these complex, relational dependencies [20].

1. Heterogeneous Graph Construction. A heterogeneous graph G = (V, E) is constructed based on the data foundations from Section 2:

- **Nodes (V):** Covered Entities (CE), Contract Pharmacies (CP), Providers (NPI), Patients (hashed ID), and Shipping Addresses.

- **Node Features:** Node features are tabular in origin, and the mapping of tabular attributes into graph learning is surveyed in [21]. Features for each node are drawn from OPAIS (e.g., Entity_Type), the enrichment data (e.g., Geospatial_Plausibility_Score), and the output of Module 1 (e.g., Anomaly_Score).

- **Edges (E):** Edges represent relationships, such as (CE)-[contracts]->(CP) (from OPAIS), (CE)-[ships_to]->(Address) (from OPAIS), (Provider)-[prescribes]->(Patient), (Patient)-[fills_at]->(CP), and (CE)-[employs]->(Provider).

2. GNN Model Specification: Inductive GraphSAGE. The framework proposes using GraphSAGE, an inductive GNN model [7]. This choice is deliberate and critical. Many GNN architectures (like Graph Convolutional Networks) are transductive, meaning they require the entire graph, including all nodes, to be present during training. This is non-viable for a dynamic, real-world system where new CEs, CPs, and patients are registered daily.

GraphSAGE, by contrast, is inductive [7]. It learns an aggregator function that generates node embeddings by sampling and aggregating features from a node's local neighborhood. This means the trained GraphSAGE model can generate fraud scores and embeddings for brand-new CEs or CPs that were not in the original training set, enabling near-real-time detection without the need to retrain the entire multi-billion-edge graph.

Expected Outcomes (GNN).

- **Node-Level Anomaly:** The GNN learns a low-dimensional embedding for "normal" CPs, providers, etc. It can then identify nodes that are structural outliers, for example a CP whose patient and provider neighborhood looks structurally different from all other "honest" CPs [22].

- **Community Detection (Collusion):** This is the GNN's most powerful application. The model can identify "collusive communities" [23], dense subgraphs of CEs, providers, and CPs that transact heavily and suspiciously with each other but are largely disconnected from the main, "honest" graph. This is the "smoking gun" for organized diversion rings.

4. Privacy-Preserving Methodologies

4.1 Rationale: The HIPAA/HITECH Imperative

The framework's reliance on private claims, prescription, and EHR data makes privacy a paramount concern, not an afterthought. This data constitutes Protected Health Information (PHI) and is strictly governed by the Health Insurance Portability and Accountability Act (HIPAA) and the HITECH Act [8]. The creation of a single, centralized "340B fraud database" containing PHI from all U.S. covered entities is legally, ethically, and logistically non-viable [8]. Therefore, Privacy-Enhancing Technologies (PETs) are a foundational requirement for this framework.

4.2 Architecture 1: Federated Learning (FL) for Collaborative Anomaly Detection

Mechanism: To train Module 1 (the Autoencoder) effectively, we propose a Federated Learning (FL) architecture [24].

Process flow:

1. A central server (e.g., managed by a regulatory body or consortium) initializes a "global model" (the AE structure).
2. This model is distributed to all participating CEs.
3. Each CE trains the model locally on its own private claims data [25]. The sensitive PHI never leaves the CE's firewall.
4. The CEs do not send data back. Instead, they send only the resulting model updates (e.g., anonymized gradients or model weights).
5. The central server aggregates these updates (e.g., via Federated Averaging) to create an improved "global model," which is then redistributed for the next round of training.

An AE trained on a single CE's "normal" data would be inherently biased. FL allows all CEs to collaboratively build a "global" model of "normal" 340B behavior [24] that is far more robust, generalized, and accurate than any single CE could build alone. This approach solves the data-silo problem without violating data sovereignty or HIPAA.

4.3 Architecture 2: Differential Privacy (DP) for Patient-Level Guarantees

Rationale: While FL is a powerful architectural defense, it is not perfectly private. Malicious actors could potentially infer information about a CE's training data by analyzing its specific model updates.

Mechanism: The FL framework must be strengthened by integrating Differential Privacy (DP) [26].

Process: During the local training step, CEs will not use standard gradient descent. Instead, they will use Differentially Private Stochastic Gradient Descent (DP-SGD) [27]. This technique introduces two key steps before the update is calculated:

1. **Gradient Clipping:** The influence of any single data point (e.g., one patient's claim) on the gradient is capped.
2. **Noise Addition:** Calibrated statistical "noise" (e.g., from a Gaussian mechanism) is added to the gradients.

This provides a formal, mathematical guarantee that an attacker, even with full access to the model updates, cannot determine whether any specific patient's data was included in the training set. This moves privacy from an "architectural" defense (FL) to a "mathematical" guarantee (DP). Implementation will require careful tuning of the privacy budget (epsilon) to balance the trade-off between privacy and model accuracy.

4.4 Synthetic Data Generation for Validation

To develop, test, and validate this complex framework without risking PHI, the use of generative models is proposed. Generative Adversarial Networks (GANs) [28] or the VAEs from Module 1 [5] can be trained to create high-fidelity synthetic healthcare claims data [29]. This synthetic data will mimic the statistical properties and distributions of the real data but will contain no link to real patient information. This allows for a safe, open environment for model development, validation, and benchmarking.

5. Implementation Challenges and Expected Outcomes

5.1 Evaluation Metrics for Severe Class Imbalance

By definition, 340B fraud is a "rare event" problem. In similar financial datasets, the fraudulent class accounts for as little as 0.17% of all transactions. In such a severely imbalanced scenario, "accuracy" is a useless and misleading metric. A model that simply flags 0% of transactions as fraudulent would be 99.8% "accurate" while providing zero detection value.

The framework's performance must be evaluated using metrics robust to this imbalance [17]:

1. **Area Under the Precision-Recall Curve (AUPRC):** This is the gold standard for imbalanced classification, as it evaluates the trade-off between precision and recall without being skewed by the large number of "true negatives".
2. **Precision@k / Recall@k:** These metrics answer the practical question for an auditor: "Of the top k providers (e.g., k = 100) flagged by the model, how many were truly high-risk?" This aligns model performance with finite auditor capacity.
3. **Area Under the Receiver Operating Characteristic (AUC-ROC):** A common metric for evaluating binary classifier performance [20].

5.2 Challenge 1: The "Black Box" Problem and Explainable AI (XAI)

A "black box" AI model that provides a fraud score with no justification is a non-starter for regulation and compliance [30]. An auditor must be given a reason why a CE, provider, or transaction was flagged.

Proposed solution (XAI):

1. **For Module 1 (AE/iForest):** The framework will implement SHAP (SHapley Additive exPlanations) [9]. For any transaction flagged as an anomaly, SHAP provides a "force plot" that quantifies which features contributed to its high anomaly score (e.g., feature 'Claim_Amount' = +3.5 to score, feature 'Procedure_Code_Mismatch' = +2.8 to score).
2. **For Module 2 (GNN):** The framework will implement GNNExplainer [10], with a fraud-specific application reported in [31]. When the GNN flags a node (e.g., a CP) as fraudulent, GNNExplainer will identify and visualize the critical subgraph and node features responsible for that decision [32]. The explanation becomes: "This pharmacy was flagged because it is linked to this set of 3 providers and 50 patients who form a suspicious, isolated community, and its own Geospatial_Plausibility_Score is low."

5.3 Challenge 2: Model Lifecycle and Concept Drift

Fraud is adversarial. Fraudsters will adapt their tactics to evade the new AI model [4]. Consequently, a model trained on historical data will see its performance decay over time as "concept drift" occurs, that is, the statistical properties of "fraud" change [30].

Proposed solution (dual-system):

1. **Drift Detection Unit:** The framework must include a monitoring component (e.g., using K-L divergence or advanced semantic shift detectors [30]) to compare the statistical distribution of new, incoming data against the original training data. A significant drift automatically triggers a model-retraining flag.
2. **Active Learning (AL) / Human-in-the-Loop (HITL):** This is the mechanism for adapting to the drift. An AL system identifies the most uncertain or novel anomalies flagged by the models. It routes only these high-value cases to a human expert auditor. The auditor's feedback ("This new pattern is indeed fraud") provides a new, high-value label. This feedback creates a continuous retraining loop, allowing the model to adapt to new fraud tactics while optimally using scarce human auditor time.

5.4 Challenge 3: Deployment and Regulatory Architecture

Given the sensitivity of PHI [8] and the high computational requirements (e.g., GPUs for GNN/AE training), the deployment model is a critical, non-trivial choice.

Proposed solution: private/hybrid cloud or on-premise.

- A **Public Cloud** (i.e., a shared, multi-tenant environment) is not suitable for this level of data sensitivity and regulatory scrutiny.

- **On-Premise:** CEs could host all data and run models entirely behind their own firewalls. This offers maximum control and security [33] but incurs high infrastructure costs and is difficult to scale.

- **Private Cloud (VPC):** This is the most viable solution. Each CE operates within a Virtual Private Cloud (VPC) on a major cloud provider (e.g., AWS, Azure). This environment is logically isolated, can be configured for HIPAA compliance, and provides the on-demand GPU resources needed for AI training [34]. This architecture also allows for the secure, encrypted communication between CE-hosted "nodes" required for the Federated Learning (FL) component.

Table III synthesizes the entire proposed framework, mapping each component to its function.

TABLE III. Model Specification and Validation Framework

Module	Model	Fraud Target	Privacy Method	XAI Method	Concept Drift Response
Module 1: Transactional	Autoencoder (AE) / VAE	Anomalous billing patterns,	Federated Learning (FL)	SHAP [9]	Retrain global model via FL

		phantom billing [3, 22]			
Module 1: Transactional	Isolation Forest (iForest)	Extreme transactional outliers	FL (for feature scaling) / Local Model	SHAP	Retrain local model
Module 2: Relational	GraphSAGE (GNN) [7]	Collusive diversion rings, structural anomalies [23, 20]	DP-SGD (via FL)	GNNExplainer [10]	Active Learning / HITL

6. Conclusion

The incumbent rule-based paradigm for 340B program oversight [4] is obsolete. Its focus on static, administrative compliance creates a "compliance paradox" where audit pass rates are high while sophisticated, networked fraud likely continues unabated.

This paper has proposed a novel framework that fundamentally shifts the oversight model from reactive compliance-checking to proactive, pattern-based detection. The framework is:

- **Hybrid:** It combines the strengths of unsupervised anomaly detection (AE/iForest) [18] for identifying entity-level outliers with the relational power of GNNs (GraphSAGE) [35] to detect network-level collusion.

- **Privacy-Preserving:** It is built on a "privacy-by-design" architecture of Federated Learning and Differential Privacy, enabling collaborative model-building without compromising HIPAA-protected data.

- **Adaptive and Interpretable:** It is designed for a real-world, adversarial environment by incorporating concept drift detection and a Human-in-the-Loop (HITL) active learning cycle. Its inclusion of XAI methods (SHAP/GNNExplainer) [9, 10] makes it auditable and explainable for regulatory bodies.

By adopting such a framework, 340B program stakeholders can move beyond the "pay-and-chase" model and develop a preventative posture, identifying and dismantling complex fraud networks before they can inflict significant financial and social harm.

Future research should focus on the integration of real-time streaming analytics for sub-second detection, the use of temporal GNNs to model the evolution of fraud networks over time, and the application of Large Language Models (LLMs) for interpreting unstructured auditor notes and semantic concept drift detection.

REFERENCES

1. Congressional Research Service, "The 340B Drug Discount Program: Litigation Topics and Trends," CRS Report R48696. [Online]. Available: https://www.congress.gov/crs-product/R48696. [Accessed: Jun. 24, 2025]
2. C. J. Courtemanche and J. Gar, "What Economists Should Know About the 340B Drug Discounting Program," National Bureau of Economic Research, Working Paper 33788, 2025. [Online]. Available: https://www.nber.org/system/files/working_papers/w33788/w33788.pdf
3. N. Prova, "Healthcare Fraud Detection Using Machine Learning," SSRN Electronic Journal, 2024. doi: 10.2139/ssrn.4892805
4. Nected, "Rules-Based Fraud Detection: 7 Strategies to Protect Your US Business." [Online]. Available: https://www.nected.ai/us/blog-us/rules-based-fraud-detection. [Accessed: Aug. 19, 2025]
5. D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," arXiv:1312.6114, 2013. [Online]. Available: https://arxiv.org/abs/1312.6114
6. F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation Forest," in Proc. 8th IEEE International Conference on Data Mining (ICDM), 2008, pp. 413-422. doi: 10.1109/ICDM.2008.17
7. W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive Representation Learning on Large Graphs," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017. arXiv:1706.02216
8. The HIPAA Journal, "What is the HITECH Act? 2025 Update." [Online]. Available: https://www.hipaajournal.com/what-is-the-hitech-act/. [Accessed: Jul. 22, 2025]

9. S. M. Lundberg and S. Lee, "A Unified Approach to Interpreting Model Predictions," in Advances in Neural Information Processing Systems (NIPS), vol. 30, 2017, pp. 4765-4774. arXiv:1705.07874
10. R. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec, "GNNExplainer: Generating Explanations for Graph Neural Networks," in Advances in Neural Information Processing Systems (NeurIPS), vol. 32, 2019. arXiv:1903.03894
11. Health Resources and Services Administration, "340B Office of Pharmacy Affairs Information System." [Online]. Available: https://www.hrsa.gov/opa/340b-opais. [Accessed: Jun. 10, 2025]
12. Health Resources and Services Administration, "Welcome to 340B OPAIS." [Online]. Available: https://340bopais.hrsa.gov/. [Accessed: Jun. 10, 2025]
13. Health Resources and Services Administration, "Daily Reports, 340B OPAIS." [Online]. Available: https://340bopais.hrsa.gov/Help/Reports/DailyReports.htm. [Accessed: Jul. 8, 2025]
14. Health Resources and Services Administration, "Frequently Asked Questions: New CE Daily Report (Excel and JSON Format), 340B OPAIS." [Online]. Available: https://340bopais.hrsa.gov/help/Reports/Frequently%20Asked%20Questions%20-%20New%20CE%20Daily%20Report.htm. [Accessed: Jul. 8, 2025]
15. Health Resources and Services Administration, "340B OPAIS User Guide for Covered Entity Users." [Online]. Available: https://340bregistration.hrsa.gov/help/CoveredEntity/Resources/PDFUserGuides/340BCoveredEntityUserGuide.pdf. [Accessed: Jun. 10, 2025]
16. Geoapify, "Address Lookup API and Address Validation/Verification API." [Online]. Available: https://www.geoapify.com/solutions/address-lookup/. [Accessed: Sep. 9, 2025]
17. Y. Chen, C. Zhao, Y. Xu, C. Nie, and Y. Zhang, "Year-over-Year Developments in Financial Fraud Detection via Deep Learning: A Systematic Literature Review," arXiv:2502.00201, 2025. [Online]. Available: https://arxiv.org/abs/2502.00201
18. R. M. Zakaria, M. M. Rahman, M. T. H. Choudhury, et al., "Detecting Financial Fraud in Real-Time Transactions Using Graph Neural Networks and Anomaly Detection Techniques," Journal of Economics, Finance and Accounting Studies, vol. 7, no. 6, 2025. doi: 10.32996/jefas.2025.7.6.1
19. Y. Yoo, J. Shin, and S. Kyeong, "Medicare Fraud Detection Using Graph Analysis: A Comparative Study of Machine Learning and Graph Neural Network Approaches," IEEE Access, vol. 11, 2023, pp. 88278-88294. doi: 10.1109/access.2023.3305962
20. D. Cheng, Y. Zou, S. Xiang, and C. Jiang, "Graph Neural Networks for Financial Fraud Detection: A Review," arXiv:2411.05815, 2024. [Online]. Available: https://arxiv.org/abs/2411.05815
21. C. Li, Y. Tsai, C. Chen, and J. C. Liao, "Graph Neural Networks for Tabular Data Learning: A Survey with Taxonomy and Directions," arXiv:2401.02143, 2024. [Online]. Available: https://arxiv.org/abs/2401.02143
22. Y. Yoo, D. Shin, D. Han, S. Kyeong, and J. Shin, "Medicare Fraud Detection Using Graph Neural Networks," in Proc. 2022 International Conference on Electrical, Computer and Energy Technologies (ICECET), 2022. doi: 10.1109/icecet55527.2022.9872963
23. L. Gomes, J. Kueck, M. Mattes, M. Spindler, and A. Zaytsev, "Collusion Detection with Graph Neural Networks," arXiv:2410.07091, 2024. [Online]. Available: https://arxiv.org/abs/2410.07091
24. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. 20th International Conference on Artificial Intelligence and Statistics (AISTATS), 2017, pp. 1273-1282. arXiv:1602.05629
25. Morgan Lewis, "AI in Healthcare: Opportunities, Enforcement Risks and False Claims, and the Need for AI-Specific Compliance," Jul. 2025. [Online]. Available: https://www.morganlewis.com/pubs/2025/07/ai-in-healthcare-opportunities-enforcement-risks-and-false-claims-and-the-need-for-ai-specific-compliance. [Accessed: Aug. 5, 2025]
26. C. Dwork, "Differential Privacy," in Automata, Languages and Programming (ICALP), Lecture Notes in Computer Science, vol. 4052, Springer, 2006, pp. 1-12. doi: 10.1007/11787006_1
27. M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep Learning with Differential Privacy," in Proc. ACM SIGSAC Conference on Computer and Communications Security (CCS), 2016, pp. 308-318. arXiv:1607.00133
28. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," arXiv:1406.2661, 2014. [Online]. Available: https://arxiv.org/abs/1406.2661
29. T. Tang, J. Yao, Y. Wang, Q. Sha, H. Feng, and Z. Xu, "Application of Deep Generative Models for Anomaly Detection in Complex Financial Transactions," arXiv:2504.15491, 2025. [Online]. Available: https://arxiv.org/abs/2504.15491
30. A. Senol, G. Agrawal, and H. Liu, "Joint Detection of Fraud and Concept Drift in Online Conversations with LLM-Assisted Judgment," arXiv:2505.07852, 2025. [Online]. Available: https://arxiv.org/abs/2505.07852
31. N. H. Dao, "Explainable Fraud Detection with GNNExplainer and Shapley Values," arXiv:2509.12262, 2025. [Online]. Available: https://arxiv.org/abs/2509.12262
32. P. Baghershahi, G. Fournier, P. Nyati, and S. Medya, "From Nodes to Narratives: Explaining Graph Neural Networks with LLMs and Graph Context," arXiv:2508.07117, 2025. [Online]. Available: https://arxiv.org/abs/2508.07117
33. iMerit, "Choosing the Right Deployment Model for Healthcare AI Annotation Platform: On-Prem vs. Hybrid vs. Private Cloud." [Online]. Available: https://imerit.net/resources/blog/choosing-the-right-deployment-model-for-healthcare-ai-annotation-platform-on-prem-vs-hybrid-vs-private-cloud/. [Accessed: Sep. 23, 2025]

34. Tamr, "Cloud AI vs. On-Premises AI: What You Need to Know." [Online]. Available: https://www.tamr.com/blog/cloud-ai-vs-onpremise-ai-what-you-need-to-know. [Accessed: Sep. 23, 2025]
35. "AI-Enabled Federated Learning System for Privacy-Preserving Health Insurance Underwriting and Fraud Detection," Technical Disclosure Commons, 2025. [Online]. Available: https://www.tdcommons.org/cgi/viewcontent.cgi?article=9052&context=dpubs_series. [Accessed: Oct. 7, 2025]