ORIGINAL RESEARCH **Open access**

A Federated Autoencoder Framework for Cross-Payer Healthcare Billing Fraud Detection

Rashid Al-Mahdi[1], Khalifa Al-Suwaidi[1*], Mariam Al-Kuwari[2]

Abstract

Healthcare billing fraud imposes major financial losses globally, costing public and private payers hundreds of billions annually. It exploits fragmented healthcare payment systems where multiple insurers process overlapping patient populations without coordination, creating blind spots that enable sophisticated cross-payer fraud schemes. Individual payers cannot detect patterns such as duplicate billing across Medicare and commercial insurers because current detection models operate within isolated organizational and regulatory boundaries. Strict privacy laws like HIPAA and GDPR further prevent sharing patient-level claims data, limiting centralized analytics. To address this, a federated anomaly detection framework is proposed in which autoencoders are trained locally at each payer without exchanging raw data. Each institution learns normal billing patterns through reconstruction-based unsupervised learning and identifies anomalies via reconstruction error. A central server aggregates encoder parameters using FedAvg, optionally with differential privacy, to build a globally informed model while preserving data locality. The resulting system enables detection of cross-payer fraud patterns, such as double billing and unbundling, that single-payer systems miss, while transmitting only model parameters through secure channels. This approach provides a privacy-preserving, scalable solution for multi-payer healthcare fraud detection under strict regulatory constraints.

Keywords Federated learning, Autoencoder, Healthcare fraud detection, Anomaly detection, Privacy-preserving machine learning, Billing fraud

*Correspondence:
Khalifa Al-Suwaidi
khalifa.suwaidi@outlook.com
[1] Department of Intelligent Healthcare Systems, Qatar University, Doha, Qatar
[2] Department of Clinical AI Analytics, Hamad Bin Khalifa University, Doha, Qatar

Introduction

Healthcare billing fraud constitutes a systemic financial burden exceeding one hundred billion dollars annually across Medicare, Medicaid, and commercial insurance programs worldwide [1]. The typology of fraudulent billing encompasses upcoding of procedure codes to inflate reimbursement, unbundling of comprehensive service packages into separately billable line items, double-billing identical procedures to multiple payers, phantom billing for services never rendered, and kickback arrangements disguised as legitimate referrals [2]. These schemes grow increasingly sophisticated as perpetrators adapt to evolving detection methodologies, leveraging their understanding of payer-specific audit thresholds and data silos to systematically extract fraudulent payments. The magnitude of this problem demands detection architectures that can operate across institutional boundaries without violating the privacy constraints that define modern healthcare data governance.

The fundamental vulnerability in current detection paradigms arises from the fragmented payer landscape, where fraud perpetrators strategically submit overlapping or contradictory claims to multiple insurers that lack the legal and technical

capacity to reconcile their respective datasets [3]. A provider can bill Medicare for a cardiac procedure while simultaneously billing a commercial payer for the identical service, knowing that neither organization possesses visibility into the other's claims adjudication. Individual payer machine learning models, regardless of their sophistication, train exclusively on the claims universe visible to that single organization and therefore remain structurally blind to cross-payer fraud patterns. The economic incentive to exploit these inter-payer data gaps continues to grow as healthcare expenditure expands across both public and private sectors.

Privacy regulations form an immovable constraint on any proposed multi-payer fraud detection solution, with HIPAA in the United States and GDPR in the European Union imposing severe penalties for unauthorized sharing of protected health information [4]. Even within less restrictive regulatory environments, competitive dynamics among commercial insurers create powerful disincentives against sharing claims data that could reveal proprietary pricing structures, network compositions, or actuarial methodologies. Centralized fraud detection systems that require pooling patient-level claims data from multiple payers remain legally and commercially infeasible despite their theoretical effectiveness at identifying cross-payer anomalies. This regulatory and competitive reality necessitates architectures that derive collaborative intelligence without data centralization.

Table 1 clarifies how the proposed framework converts the core barriers of cross-payer healthcare fraud detection into specific architectural design responses.

Table 1. Analytical alignment between fraud-detection barriers and federated autoencoder design responses

Fraud-detection barrier	Why the barrier matters in cross-payer fraud	Framework design response	Analytical contribution
Cross-payer data fragmentation	Fraud schemes can be distributed across insurers, making them invisible within any single payer’s claims environment	Federated learning enables multiple payers to train a shared detection representation without pooling claims	Converts institutional fragmentation from a detection weakness into a distributed learning structure
Scarcity of labelled fraud examples	Confirmed fraud labels are rare, delayed, costly, and biased toward previously detected schemes	Autoencoders learn normal billing manifolds through reconstruction loss rather than supervised fraud labels	Enables anomaly discovery under weak-label and low-label conditions
Privacy and regulatory constraints	HIPAA, GDPR, and commercial confidentiality prevent centralized claims aggregation	Only encoder parameters are exchanged; raw claims and local decoders remain within payer boundaries	Aligns fraud analytics with legal and institutional data-governance constraints
Heterogeneous payer populations	Payers differ in demographics, provider networks, benefit structures, and regional practice patterns	Weighted FedAvg incorporates payer-specific information while synthesizing consortium-level billing regularities	Supports model generalization while preserving sensitivity to distributed payer variation
Adaptive fraud behavior	Fraud perpetrators modify billing strategies in response to detection rules and audits	Reconstruction-error scoring can surface deviations from normative billing behavior beyond predefined rules	Improves detection of novel or evolving fraud schemes
Investigation-capacity limits	Fraud units cannot manually review all flagged claims at equal	Ranked anomaly queues prioritize high-error and cross-	Connects model output to operational triage and

	priority	payer-confirmed claims	investigator efficiency

This paper proposes a federated autoencoder framework that enables multiple healthcare payers to collaboratively train anomaly detection models for billing fraud without exchanging any patient-level claims data. Each participating payer trains a local autoencoder on its own claims repository and contributes only encoder parameters to a federated aggregation process that synthesizes cross-payer billing pattern knowledge into a shared model. The framework addresses the fundamental tension between data locality requirements and the need for multi-institutional pattern detection in healthcare fraud analytics [5].

Background

Healthcare billing fraud schemes

Healthcare billing fraud manifests through distinct typologies that exploit the structural complexity of medical coding and reimbursement systems [6]. Upcoding involves the deliberate misrepresentation of a procedure or diagnosis code to one that commands a higher reimbursement rate, such as billing a comprehensive evaluation when only a limited examination occurred. Unbundling systematically separates procedure components that reimbursement rules require to be billed as a single comprehensive code, artificially inflating total claim value through fragmentation [7]. Double-billing represents the submission of identical claims to multiple payers, while phantom billing invoices for services, equipment, or prescriptions that were never actually provided to the patient. These schemes share a common characteristic: they are designed to appear legitimate when examined in isolation by a single payer’s adjudication system.

Current fraud detection limitations

Existing healthcare fraud detection methodologies operate predominantly within single-payer data environments, employing rule-based heuristics, supervised machine learning classifiers, and anomaly detection algorithms trained on organization-specific claims histories [8]. Rule-based systems flag transactions that violate predefined parameters such as maximum daily service volumes or incompatible procedure-diagnosis pairings, but fail to detect novel or evolving fraud vectors that fall outside encoded logic. Supervised models require labelled fraud examples that are scarce, expensive to produce through manual investigation, and biased toward previously identified scheme patterns. Even the most advanced single-payer detection systems cannot identify a claim that appears legitimate within that payer's data context but replicates an identical service already billed to another insurer for the same patient on the same date [9].

Autoencoders for anomaly detection

Autoencoders provide an unsupervised deep learning paradigm particularly suited to fraud detection scenarios where labelled fraudulent examples are rare and skewed toward historical patterns [10]. An autoencoder consists of an encoder network that compresses input data into a lower-dimensional latent representation and a decoder network that reconstructs the original input from this compressed encoding. The model is trained to minimize reconstruction error on a dataset assumed to represent normal behaviour, learning the manifold structure of legitimate transactions. When presented with anomalous inputs that deviate from learned normal patterns, the autoencoder produces elevated reconstruction error that serves as a natural anomaly score without requiring explicit fraud labels during training [11]. This reconstruction-based anomaly detection approach has demonstrated effectiveness in healthcare claims contexts where the vast majority of transactions represent legitimate billing activity.

Federated learning

Federated learning establishes a distributed machine learning paradigm where multiple data-holding parties collaboratively train a shared model without centralizing their raw data [12]. The foundational Federated Averaging algorithm proceeds through alternating rounds of local model training and global parameter aggregation, wherein each

participating client computes gradient updates on its private dataset and transmits only the model parameters to a central server. The server performs a weighted average of received parameters, typically proportional to each client's dataset size, to produce an updated global model that synthesizes knowledge across all participants [13]. This architecture provides fundamental privacy guarantees by ensuring raw data never leaves its source institution, with differential privacy mechanisms offering formal bounds on information leakage through parameter transmissions [14]. Multiple variations have emerged to address heterogeneous data distributions, communication efficiency constraints, and adversarial robustness requirements across diverse deployment contexts.

Framework Overview

High-level architecture

The proposed framework establishes a federated learning topology where multiple healthcare payer organizations function as independent clients, each operating a local autoencoder trained exclusively on its own adjudicated claims database [15]. A central aggregation server orchestrates the federated training process without ever receiving or processing patient-level claims data, coordinating secure communication channels and parameter synchronization schedules. Each payer trains its local autoencoder on historical claims assumed to represent normal billing patterns, then periodically transmits encoder network weights to the aggregation server for global model synthesis. The resulting federated autoencoder, distributed back to all participating payers, enables each organization to compute anomaly scores on incoming claims using a model informed by billing patterns from the entire payer consortium, thereby exposing cross-payer fraud patterns that individual models would inevitably miss.

Figure 1 presents the proposed federated autoencoder architecture for detecting cross-payer healthcare billing fraud while preserving payer-level data locality and patient privacy.

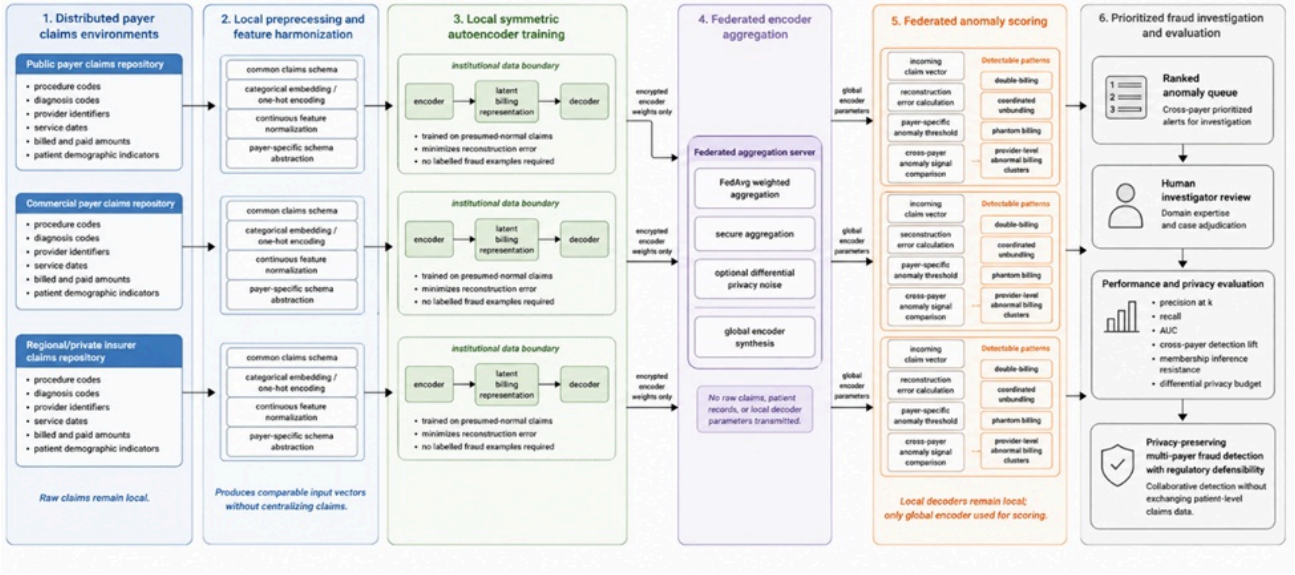

Figure 1. Federated Autoencoder Architecture for Privacy-Preserving Cross-Payer Healthcare Billing Fraud Detection

Core assumptions

The framework operates under several foundational assumptions necessary for multi-payer collaboration in healthcare fraud analytics [16]. Participating payers must agree upon and implement a standardized claims data schema encompassing common fields such as procedure codes, diagnosis codes, provider identifiers, billed amounts, and service dates, enabling consistent feature encoding across institutions with divergent native data formats. A secure

communication infrastructure must exist between each payer and the aggregation server, supporting encrypted parameter transmission and authenticated client identity verification. The framework further assumes that participating payers possess sufficient computational resources to train deep autoencoder models on their local claims repositories and that a mutually acceptable governance structure defines participation terms, aggregation schedules, and model usage rights.

Design principles

Three core design principles govern the framework architecture, balancing privacy preservation with detection effectiveness across the multi-payer consortium [17]. Privacy preservation constitutes the non-negotiable foundation, requiring that no patient-level claims data traverse organizational boundaries, with model parameter transmission representing the sole inter-payer information exchange. Scalability must accommodate growing payer participation and expanding claims volumes without proportional increases in communication overhead or aggregation complexity, ensuring practical deployment across large and heterogeneous payer networks. Robustness to heterogeneous claims distributions across payers, arising from different patient demographics, regional practice patterns, and benefit structures, requires architectural choices that maintain anomaly detection performance despite non-identically distributed training data across federation participants.

Local Autoencoder Training

Claims data encoding

Each payer transforms its proprietary claims records into a standardized feature representation suitable for autoencoder training, encoding both categorical attributes and continuous values into a unified numerical vector [18]. Categorical features including provider National Provider Identifier, patient demographic indicators, procedure codes, and diagnosis codes undergo embedding layer transformation or one-hot encoding depending on cardinality and domain semantics. Continuous features such as billed amount, paid amount, service date temporal encoding, and procedure quantity measures are normalized to consistent numerical ranges to prevent gradient dominance during neural network optimization. The resulting feature vectors capture the multidimensional structure of healthcare billing transactions while abstracting away payer-specific data schema variations through the common encoding specification agreed upon by federation participants.

Autoencoder architecture

The local autoencoder architecture employs symmetric encoder and decoder networks connected through a bottleneck latent representation that captures the essential manifold of legitimate billing patterns [19]. The encoder progressively compresses the input claims feature vector through a series of fully connected layers with decreasing dimensionality, applying rectified linear unit activations at intermediate layers to introduce non-linearity while mitigating vanishing gradient concerns. The decoder mirrors this structure in reverse, expanding from the latent representation through layers of increasing dimensionality to reconstruct the original input vector at the output layer, with sigmoid activation applied at the final layer for features normalized to the unit interval. This symmetric design ensures that the reconstruction objective drives the latent space to capture the principal modes of variation present in the payer's normal claims distribution.

Training objective

Local autoencoder training minimizes mean squared reconstruction error between the input claims feature vector and the decoder output, computed across a training corpus assumed to represent predominantly legitimate billing transactions [20]. The reconstruction loss provides an unsupervised training signal that requires no fraud labels, driving the autoencoder to learn a compressed representation that preserves the information necessary to reproduce normal claims while discarding idiosyncratic noise. Each payer implements early stopping based on validation set reconstruction error to prevent overfitting, and hyperparameter optimization across network depth, latent dimensionality, and learning rate configurations identifies architectures appropriate to the payer's claims volume and feature complexity. The assumption

that training data consists overwhelmingly of legitimate claims enables the autoencoder to learn normal billing patterns against which future claims can be evaluated for anomalous deviation.

Federated Aggregation

Encoder aggregation

The federated aggregation process synthesizes a globally-informed encoder by weighted averaging of locally-trained encoder parameters transmitted from each participating payer to the central server [21]. The aggregation server computes the global encoder weights as a weighted sum of local encoder parameters, where the weight assigned to each payer's contribution corresponds to the proportion of total consortium training examples that payer contributed to the current round. This FedAvg weighting scheme ensures that payers with larger claims volumes exert proportionally greater influence on the global model while still incorporating pattern information from smaller payers whose specialized patient populations may expose distinct fraud vectors. The aggregated encoder is then distributed back to all payers, each of which pairs it with their locally-trained decoder for subsequent anomaly scoring operations without sharing decoder parameters that might encode payer-specific reimbursement characteristics.

Privacy enhancements

Additional privacy protection mechanisms supplement the baseline federated architecture to provide formal guarantees against information leakage through parameter transmissions [22]. Differential privacy is incorporated by adding calibrated Gaussian noise to encoder gradient updates prior to transmission, with the privacy budget epsilon controlling the trade-off between model utility and privacy guarantees, such that smaller epsilon values provide stronger protection at the potential cost of reduced anomaly detection accuracy. Secure aggregation protocols employing cryptographic techniques such as secure multi-party computation ensure that the central server can compute the weighted average of encrypted parameter vectors without ever inspecting individual payer contributions in plaintext. These layered privacy protections address regulatory requirements under HIPAA and GDPR while maintaining the collaborative learning capability essential for cross-payer fraud pattern detection.

Anomaly Detection and Scoring

Anomaly score calculation

Upon deployment of the federated autoencoder, each payer computes anomaly scores for incoming claims by measuring the mean squared reconstruction error between the input feature vector and the decoder output [23]. The reconstruction error quantifies the degree to which a given claim deviates from the normal billing patterns encoded in the federated model's latent representation, with legitimate claims expected to reconstruct accurately and fraudulent or anomalous claims producing substantially higher error magnitudes. A detection threshold is established by analyzing the distribution of reconstruction errors across a held-out validation set of claims presumed normal, with the 99th percentile of this error distribution serving as an initial operational threshold that can be calibrated based on payer-specific investigation capacity and fraud prevalence estimates. Claims exceeding this threshold are flagged for further scrutiny, providing an unsupervised mechanism that identifies both known fraud typologies and novel schemes that manifest as deviations from normative billing behaviour.

Cross-payer anomaly scoring

The federated architecture enables a novel cross-payer scoring capability where claims submitted to multiple payers generate anomaly signals that can be aggregated across the consortium to strengthen detection confidence [24]. When a provider submits substantially similar claims to multiple participating payers, each payer independently computes a reconstruction error score using the federated autoencoder, and these scores can be compared or aggregated through

the central server without exposing the underlying patient data. Elevated anomaly scores for the same provider, service date, and procedure code combination appearing simultaneously across multiple payers provide strong evidence of potential double-billing or coordinated fraud schemes that single-payer models would evaluate in isolation. This cross-payer scoring mechanism represents the primary detection advantage of the federated framework, directly addressing the multi-payer blind spot that current detection systems cannot resolve.

Fraud Investigation Workflow

Triage and prioritization

Flagged claims enter a structured triage workflow that ranks anomalies by severity and investigation priority, ensuring efficient allocation of limited fraud investigation resources [25]. Claims are sorted by their anomaly score magnitude, with the highest-scoring transactions receiving immediate investigative attention, while cross-payer flags where multiple payers independently generate elevated scores for the same provider-service combination receive automatic priority escalation. The triage system can incorporate additional business rules such as claim dollar amount thresholds and provider historical fraud rates to further refine prioritization, combining the data-driven anomaly signal with domain-specific investigative heuristics. This layered prioritization ensures that investigators focus on the highest-probability and highest-impact fraud cases rather than being overwhelmed by uniformly processed anomaly alerts.

Investigation feedback loop

The framework incorporates a closed-loop feedback mechanism where confirmed fraud determinations from human investigations are systematically reintegrated into the model improvement cycle [26]. When investigators validate that a flagged claim represents actual fraud, this labelled example can be used to refine local autoencoder training by either excluding confirmed fraudulent claims from the normal training corpus or incorporating the fraud label into a semi-supervised training objective. Investigation outcomes that determine flagged claims to be legitimate false positives similarly provide valuable signal, indicating billing patterns that the autoencoder incorrectly identifies as anomalous and suggesting feature engineering or threshold adjustments to reduce future false alarm rates. This continuous feedback loop enables the detection system to adapt to evolving fraud patterns and improve specificity over successive investigation cycles.

Evaluation Strategy

Table 2 defines a validation logic for assessing whether the federated autoencoder framework delivers detection gains without undermining privacy, robustness, or operational feasibility.

Table 2. Evaluation logic for validating technical performance, operational usefulness, and privacy preservation

Evaluation dimension	Core question	Suggested metric or test	Interpretation of strong performance
Local anomaly discrimination	Can each payer identify claims that deviate from its learned normal billing distribution?	Reconstruction-error distribution, AUC, recall at fixed threshold	The model separates anomalous claims from presumed-normal billing patterns within each payer
Investigation efficiency	Does the model improve the usefulness of fraud review queues?	Precision at k, fraud yield per investigator hour, false-positive burden	Top-ranked claims contain a higher concentration of confirmed fraud than conventional rule-based queues

Cross-payer detection lift	Does federated learning detect fraud that single-payer models miss?	Detection lift over independent payer baselines, stratified by fraud type	Federated scoring improves detection of double-billing, coordinated unbundling, and provider-level cross-payer schemes
Robustness to payer heterogeneity	Does performance remain stable across payer size, population mix, and benefit structure?	Subgroup performance by payer type, claim volume, geography, and provider network	Detection performance does not collapse for smaller or demographically distinct payers
Privacy leakage resistance	Do exchanged parameters reveal sensitive patient or payer information?	Membership inference attack success, mutual information leakage, differential privacy epsilon tracking	Parameter sharing provides measurable privacy protection and supports regulatory defensibility
Governance and deployment feasibility	Can competing payers sustain a shared detection system?	Participation compliance, aggregation reliability, communication overhead, auditability	The system operates within practical institutional, technical, and legal constraints

Detection metrics

Rigorous evaluation of the federated fraud detection framework requires metrics that capture both the accuracy of anomaly ranking and the operational efficiency of the investigation workflow [27]. Precision at rank k measures the fraction of the top k highest-scoring claims that correspond to confirmed fraud, directly quantifying the efficiency of investigator time allocation when reviewing prioritized alerts. Recall metrics assess the proportion of total fraud captured within various anomaly score thresholds, while the area under the receiver operating characteristic curve provides a threshold-independent measure of overall detection discrimination when labelled evaluation data is available. These metrics should be computed at both the individual payer level and across the entire consortium to characterize detection performance across heterogeneous claims distributions.

Cross-payer detection lift

The principal performance claim of the federated framework must be evaluated by directly comparing cross-payer fraud detection rates against the baseline of independent single-payer models operating without parameter sharing [28]. Detection lift is quantified by measuring the proportion of known cross-payer fraud cases that the federated model correctly flags compared to the single-payer baseline, with particular attention to double-billing and coordinated unbundling schemes that span multiple insurers. This comparative analysis should stratify results by fraud scheme type, claim dollar amount, and provider characteristics to identify the specific scenarios where federated collaboration provides the greatest detection advantage. Statistical significance testing on the observed lift validates that improvements derive from the multi-payer architecture rather than random variation in detection performance.

Privacy metrics

Privacy preservation claims require quantitative evaluation through established information security and privacy auditing methodologies adapted to the federated learning context [29]. Membership inference attack resistance measures the extent to which an adversary observing model parameters can determine whether a specific patient's claims were included in the training data, providing an empirical privacy leakage assessment complementary to formal differential privacy guarantees. The differential privacy epsilon budget is tracked across aggregation rounds to verify compliance with predetermined privacy loss limits, while information leakage measurements quantify the mutual information between transmitted parameter updates and sensitive attributes of the underlying training data. These privacy metrics establish

verifiable bounds on the information disclosed through the federated training process, supporting regulatory compliance demonstrations to data protection authorities and institutional review boards.

Limitations

Technical limitations

The framework relies on the foundational assumption that the training corpus consists predominantly of legitimate claims, an assumption that may be violated in payers with high fraud penetration rates, introducing label noise that degrades the autoencoder's learned representation of normal billing patterns [12]. Autoencoder architectures are optimized for input reconstruction fidelity rather than discriminative fraud separation, meaning that subtle fraudulent claims lying close to the legitimate data manifold may reconstruct accurately and evade detection despite representing genuine fraud. Adversarial fraud perpetrators with sufficient technical sophistication could conceivably craft claims designed to produce low reconstruction errors against the federated model by learning to mimic the latent representations of legitimate transactions, representing a limitation inherent to unsupervised anomaly detection approaches.

Operational limitations

Effective deployment requires sustained cooperation among competing payer organizations that must agree on data standardization, aggregation protocols, and shared governance structures, presenting organizational challenges that may exceed the technical complexity of the federated system itself [3]. False positive anomalies that survive triage prioritization incur investigation costs and may strain payer-provider relationships when legitimate claims face audit scrutiny, necessitating careful threshold calibration that balances detection sensitivity against operational practicality. Fraud perpetrators represent adaptive adversaries who modify their schemes in response to detection capabilities, introducing concept drift in claims distributions that requires continuous model retraining through the investigation feedback loop to maintain detection effectiveness over time.

Conclusion

The federated autoencoder framework establishes a viable architectural pathway toward privacy-preserving, multi-payer healthcare billing fraud detection that reconciles the competing demands of cross-institutional pattern recognition and patient data protection. By training autoencoders locally at each payer and aggregating only encoder parameters through a federated server, the framework enables collaborative anomaly detection without any raw claims data leaving its originating institution. The architecture addresses the critical blind spot that current single-payer detection systems face when confronting fraud schemes strategically distributed across multiple insurers.

The framework offers several distinctive advantages over existing approaches, including the ability to detect cross-payer fraud patterns such as double-billing and coordinated unbundling that no individual payer can identify independently. The unsupervised autoencoder training paradigm eliminates the requirement for labelled fraud examples, which are chronically scarce in healthcare fraud analytics, instead learning the manifold of legitimate billing transactions from which anomalies naturally deviate. The layered privacy protections, incorporating differential privacy and secure aggregation protocols, provide regulatory defensibility under HIPAA and GDPR frameworks that prohibit patient data centralization.

Important limitations must be acknowledged, including the assumption that training data consists predominantly of legitimate claims, the operational challenge of securing sustained cooperation among competing payer organizations, and the inherent tension between detection sensitivity and false positive investigation burden. Fraud perpetrators represent adaptive adversaries whose evolving schemes necessitate continuous model refinement through investigation feedback loops, and the autoencoder's reconstruction objective may fail to discriminate subtle fraudulent claims that lie close to the normal data manifold.

Implementation on multi-payer claims databases, such as the Centers for Medicare and Medicaid Services research datasets in conjunction with commercial claims consortiums, represents the essential next step toward validating the framework's detection effectiveness in operational healthcare environments. Such real-world deployment would generate the empirical evidence necessary to quantify cross-payer detection lift, characterize privacy-utility trade-offs under various differential privacy budgets, and establish best practices for federated healthcare fraud analytics at scale.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 01 Dec 2024 Revised: 18 Dec 2024 Accepted: 17 Jan 2025
Published online: 20 July 2025

References

Yang Q, Liu Y, Chen T, Tong Y. Federated machine learning: concept and applications. ACM Trans Intell Syst Technol. 2019;10(2):1–9.

Bauder R, Khoshgoftaar T. A survey of Medicare data processing and integration for fraud detection. In: 2018 IEEE International Conference on Information Reuse and Integration (IRI). IEEE; 2018:9–14.

Johnson JM, Khoshgoftaar TM. Medicare fraud detection using neural networks. J Big Data. 2019;6(1):63.

Rieke N, Hancox J, Li W, Milletarì F, Roth HR, Albarqouni S, et al. The future of digital health with federated learning. NPJ Digit Med. 2020;3:119.

https://doi.org/10.1038/s41746-020-00323-1.

Xu J, Glicksberg BS, Su C, Walker P, Bian J, Wang F. Federated Learning for Healthcare Informatics. J Healthc Inform Res. 2021;5(1):1-19.
https://doi.org/10.1007/s41666-020-00082-4.

Herland M, Bauder RA, Khoshgoftaar TM. Approaches for identifying US Medicare fraud in provider claims data. Health Care Manag Sci. 2020;23(1):2–19.

Hancock JT, Bauder RA, Wang H, Khoshgoftaar TM. Explainable machine learning models for Medicare fraud detection. J Big Data. 2023;10(1):154.

Nabrawi E, Alanazi A. Fraud detection in healthcare insurance claims using machine learning. Risks. 2023;11(9):160.

du Preez A, Bhattacharya S, Beling P, Bowen E. Fraud detection in healthcare claims using machine learning: a systematic review. Artif Intell Med. 2025;160:103061.

Siddalingappa R, Kanagaraj S. Anomaly detection on medical images using autoencoder and CNN. Int J Adv Comput Sci Appl. 2021;12(7).

Novoa-Paradela D, Fontenla-Romero O, Guijarro-Berdiñas B. Fast deep autoencoder for federated learning. Pattern Recognit. 2023;143:109805.

McMahan B, Moore E, Ramage D. Communication-efficient learning of deep networks from decentralized data. In: AISTATS. PMLR; 2017:1273–82.

Kairouz P, McMahan HB. Advances and open problems in federated learning. Found Trends Mach Learn. 2021;14(1–2):1–210.

Lyu L, Yu H, Yang Q. Threats to federated learning: a survey. arXiv preprint. 2020:arXiv:2003.02133.

Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, et al. Practical secure aggregation for privacy-preserving machine learning. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security; 2017. p. 1175-91.

Guan H, Yap PT, Bozoki A, Liu M. Federated learning for medical image analysis: a survey. Pattern Recognit. 2024;151:110424.

Mothukuri V, Parizi RM, Pouriyeh S, Huang Y, Dehghantanha A, Srivastava G. A survey on security and privacy of federated learning. Future Gener Comput Syst. 2021;115:619–40.

Dang TK, Lan X, Weng J, Feng M. Federated learning for electronic health records. ACM Trans Intell Syst Technol. 2022;13(5):1–7.

Cai Y, Chen H, Cheng KT. Rethinking autoencoders for medical anomaly detection. In: MICCAI. Springer; 2024:544–54.

Nawaz A, Khan SS, Ahmad A. Ensemble of autoencoders for anomaly detection in biomedical data. IEEE Access. 2024;12:17273–89.

Li S, Cheng Y, Wang W, Liu Y, Chen T. Learning to detect malicious clients for robust federated learning. arXiv preprint. 2020:arXiv:2002.00211.

Nawaz A, Khan SS, Ahmad A, Ghenimi N, Ahmed LA. Federated anomaly detection for neonatal health. Inform Med Unlocked. 2025:101713.

Kumaraswamy N, Markey MK, Ekin T, Barner JC, Rascati K. Healthcare fraud data mining methods: a look back and look ahead. Perspect Health Inf Manag. 2022;19(1).

Yoo Y, Shin D, Han D, Kyeong S, Shin J. Medicare fraud detection using graph neural networks. In: ICECET. IEEE; 2022:1–5.

Muhammad R, Tbaishat D, Nazir A, Yacoub S, AbdulRazek M, El-Enen MA, et al. Fraud detection and explanation in medical claims using GNN architectures. Sci Rep. 2025;15(1).

Hong B, Lu P, Xu H, Lu J, Lin K, Yang F. Health insurance fraud detection based on heterogeneous graph learning. Heliyon. 2024;10(9).

Alam MS, Tiwari RK, Pandey V. Healthcare billing fraud detection using ML and homomorphic encryption. In: ICRTCST. IEEE; 2022:181–6.

Vepakomma P, Gupta O, Swedish T, Raskar R. Split learning for health: distributed deep learning without sharing raw data. arXiv preprint. 2018:arXiv:1812.00564.

Ünal C, Erbuğa GS. Detection and prevention of medical fraud using machine learning. Acta Infologica. 2024;8(2):100–17.