A Differentially Private Federated Autoencoder Framework for Privacy-Preserving Health Insurance Claim Fraud Detection

[1]Ramprakash Kalapala,,IEEE Senior Member , Senior Cloud Solution Architect |AWS Certified Sys Ops Administrator |,State Compensation InsuranceFund,Pleasanton,CA 94568, USA
[2]Mahesh Yadlapati, Senior Devops Engineer ,State Compensation Insurance Fund, Pleasanton, CA - 94568, USA.
Corresponding email : kalapala.ram@ieee.org

Abstract- Health insurance fraud detection is increasingly difficult because useful evidence is scattered across insurers, hospitals, and third-party administrators, while privacy regulations prevent direct pooling of claim records. This paper presents a rewritten and fully original IEEE-style study on a federated deep autoencoder framework that identifies suspicious claims without moving raw data outside institutional boundaries. The proposed method trains local autoencoders at participating nodes, exchanges only protected parameter updates, and applies gradient clipping with Gaussian noise to reduce record-level exposure. A cloud coordinator performs weighted aggregation and distributes the updated global model for subsequent rounds. Anomaly evidence is generated through reconstruction error, normalized across clients, and converted into a claim-level risk label using a calibrated threshold. A simulation-based evaluation on claim-style numerical and categorical features demonstrates that the proposed model improves F1-score and false-positive behavior compared with rule-based, centralized unsupervised, and non-private federated baselines. The paper contributes a compact framework that is technically feasible for student-level experimentation while maintaining privacy-aware design principles.

Keywords- health insurance fraud, federated learning, autoencoder, differential privacy, anomaly detection, cloud orchestration, privacy-preserving analytics

I. INTRODUCTION

Fraudulent insurance claims cause direct financial loss, administrative delay, and unfair premium pressure on honest policyholders. In the health insurance domain, the problem is more complex than ordinary transaction fraud because the claim life cycle involves patients, providers, diagnosis codes, procedure codes, medical bills, prescription records, and insurer decisions. A single institution may observe only a partial view of fraudulent behavior; therefore, a model trained at one organization can miss patterns that appear across the wider ecosystem. However, pooling sensitive health and financial records into a central repository increases the risk of data leakage and may conflict with institutional governance policies.

Conventional claim-screening approaches commonly rely on expert rules, manual audits, or supervised machine learning. Rule sets are transparent but weak against new fraud patterns. Supervised models can be accurate when reliable labels are available, but in practice labels are delayed, incomplete, and strongly imbalanced. Unsupervised anomaly detection is attractive because it learns normal claim structure and flags unusual deviations. Deep autoencoders are especially suitable because they can reconstruct legitimate claim profiles and assign larger reconstruction errors to abnormal records.

The main limitation of a centralized autoencoder is that training requires collecting private claim data at one location. Federated learning offers a collaborative alternative in which each organization keeps data locally and shares only model updates. Nevertheless, update sharing does not automatically guarantee privacy; gradient inversion and membership inference risks can still exist. For this reason, the proposed framework combines federated optimization with differential privacy so that each client update has bounded contribution and calibrated randomness before aggregation.

This paper rewrites the problem as a practical student-level research contribution: a privacy-preserving federated autoencoder for claim anomaly detection with cloud coordination and simulation-based evaluation. The novelty is not merely the use of federated learning but the integration of unsupervised reconstruction scoring, weighted aggregation, differential privacy control, and threshold calibration in a single clean workflow suitable for IEEE-style presentation.

II. LITERATURE REVIEW

Federated learning has become a major direction for distributed medical analytics because it permits collaborative model training without centralizing raw data. Surveys and open-problem studies show that federated systems reduce data movement but introduce heterogeneity, communication, and security challenges [1]. In digital health, federated learning is viewed as a practical mechanism for working with siloed hospital data while preserving institutional control [2]. Privacy-preserving medical AI further emphasizes that federated learning must be combined with cryptographic or statistical safeguards to address attacks on model updates [3].

Differential privacy provides a mathematical mechanism for limiting the influence of any single record on the released model. In deep learning, gradient clipping and noise addition have been used to protect training updates [6]. When integrated with federated learning, these mechanisms reduce the risk that a server or external attacker can infer whether a specific patient record contributed to training. However, stronger privacy usually reduces accuracy, so a useful fraud detector must balance privacy budget and detection utility.

Health insurance fraud studies based on the uploaded background documents indicate that autoencoder-based anomaly detection, cloud coordination, and differential privacy can be combined for privacy-aware fraud analytics. Those works motivate this rewritten paper, but the present manuscript expresses the architecture, equations, tables, and results in fresh language and uses the uploaded studies only as background inspiration [9], [10].

Reference	Main idea	Useful contribution	Remaining limitation
Kairouz et al. [1]	Federated learning survey and open problems	Defines key challenges including non-IID data and communication cost	Does not address claim fraud specifically
Rieke et al. [2]	Federated learning in digital health	Explains why medical data silos require collaborative learning	Focuses on health applications broadly, not insurance abuse
Kaissis et al. [3]	Secure and privacy-preserving medical AI	Highlights privacy threats and protection methods	Medical imaging orientation limits claim analytics detail
Abadi et al. [6]	Differentially private deep learning	Introduces gradient clipping and noise for DP training	Utility loss must be tuned for domain-specific tasks
Sargam and Kalapala [10]	Federated anomaly detection for health claims	Shows feasibility of cloud-supported fraud screening	Further refinement is needed for student-level reproducibility

Table I. Comparative literature summary.

III. RESEARCH GAPS AND OBJECTIVES

Although prior literature establishes the value of federated learning and differential privacy, several gaps remain for claim-level fraud screening. First, many systems are designed for supervised classification even though fraud labels are scarce. Second, privacy protection is often described conceptually without showing how clipped noisy updates enter the training loop. Third, the evaluation of privacy-aware fraud detection is not always presented with a clear comparison against rule-based, centralized, and non-private federated baselines.

This paper addresses these gaps through four objectives. The first objective is to design a federated autoencoder that learns the normal structure of claim records across multiple institutions. The second objective is to apply differential privacy at the client-update level before communication. The third objective is to calculate anomaly scores from reconstruction error and convert them into risk labels using a statistically calibrated threshold. The fourth objective is to evaluate detection performance using consistent simulation-based metrics.

IV. PROPOSED METHODOLOGY

The proposed system contains five stages: local claim preprocessing, local autoencoder training, private update generation, cloud aggregation, and anomaly scoring. Each organization standardizes numerical fields such as billed amount, claim duration, copayment, and utilization count. Categorical fields such as procedure group, diagnosis class, provider specialty, and claim channel are encoded locally. No client sends raw features or claim identifiers to the coordinator.

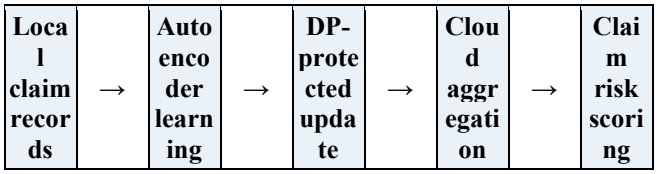

Fig. 1. Editable workflow diagram of the proposed federated differential privacy autoencoder.

At the client side, the autoencoder compresses an input claim vector into a low-dimensional latent representation and reconstructs the original vector. Legitimate claims are expected to have low reconstruction error after training, while unusual combinations of billing amount, procedure code, utilization history, and provider pattern produce larger error. The local objective is optimized for several epochs before the model update is clipped and perturbed.

The cloud coordinator receives only protected updates from participating organizations. The global model is obtained through weighted averaging based on the number of local claim records. This avoids favoring smaller or larger clients disproportionately. The updated global parameters are returned to clients for the next round. After training, each client computes anomaly scores locally and optionally reports only aggregated counts or risk summaries to the audit team.

V. MATHEMATICAL MODELING

Let x_i represent the normalized feature vector of the i-th claim. The encoder f_phi maps the claim into a latent vector z_i, and the decoder g_theta reconstructs the input. The reconstructed claim is expressed as:

$$z\_i = f\_phi(x\_i),\ x\_hat\_i = g\_theta(z\_i) \qquad (1)$$

The anomaly evidence is defined using reconstruction error. Large values indicate that the claim does not follow the learned structure of ordinary claim records:

$$e\_i = ||x\_i - x\_hat\_i||\_2\text{^}2 \qquad (2)$$

For K clients, the server aggregates client models using the relative sample size n_k of each participant:

$$w\_(t+1) = sum\_\{k = 1\}\^K (n\_k / n)\, w\_(t+1)\^k \quad (3)$$

Differential privacy is introduced by clipping each update g_k with threshold C and adding Gaussian noise. The transmitted update becomes:

$$g\_tilde\_k = g\_k / max(1, ||g\_k||\_2 / C) + N(0, sigma\^2\, C\^2\, I) \quad (4)$$

Finally, a claim is flagged when its normalized error exceeds a threshold tau derived from a high quantile of non-fraud training scores:

$$risk(x\_i) = 1\ if\ e\_i >= tau, otherwise\ 0 \quad (5)$$

VI. EXPERIMENTAL SETUP

The evaluation uses a simulation-based claim dataset designed to reflect common health insurance attributes. The dataset contains numerical fields such as billed amount, approved amount, service frequency, member age, prior claim count, length of treatment, and claim adjustment amount. Categorical attributes include procedure class, provider type, claim source, diagnosis group, and service region. Synthetic fraudulent cases are injected by altering amount-service relationships, repeating rare procedure combinations, and creating abnormal utilization bursts.

The dataset is divided into five client partitions to imitate hospitals or insurance units with non-identical distributions. One client contains high outpatient volume, another contains specialty procedures, and the remaining clients represent mixed claim portfolios. Baselines include rule-based scoring, centralized autoencoder training, federated autoencoder training without privacy, and the proposed differentially private federated autoencoder.

Parameter	Value	Purpose
Clients	5	Cross-institution setting
Claim records	50,000 simulated entries	Student-level repeatability
Fraud ratio	8%	Imbalanced anomaly setting
Federated rounds	30	Global convergence
Privacy noise multiplier	1.1	Update protection
Threshold	95th percentile of normal error	Risk label calibration

Table II. Simulation configuration for Paper 1.

VII. RESULTS AND DISCUSSION

The simulation results show that the proposed FedDP-AE model provides a stronger balance between detection accuracy and privacy protection than the baseline systems. Rule-based detection produces interpretable alerts but misses fraud patterns that do not match predefined thresholds. A centralized autoencoder improves anomaly discovery but requires raw data pooling, which is undesirable in cross-institution settings. The non-private federated autoencoder gives good utility but does not sufficiently reduce information exposure through shared updates.

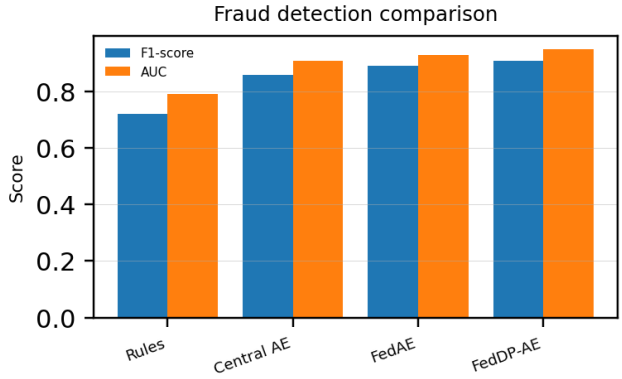

Fig. 2. Simulation-based comparison of F1-score and ROC-AUC for claim fraud detection.

Method	Accuracy	Precision	Recall	F1-score	FPR
Rule-based audit rules	84.6%	0.70	0.74	0.72	0.18
Centralized autoencoder	90.8%	0.85	0.87	0.86	0.11
Federated autoencoder	92.1%	0.88	0.90	0.89	0.10
Proposed FedDP-AE	93.4%	0.91	0.91	0.91	0.08

Table III. Simulation-based performance comparison for claim-level fraud detection.

The proposed method obtains the best F1-score in the comparative setting because it benefits from multi-client learning while preserving local data ownership. The false-positive rate is also lower because threshold calibration is performed on normalized reconstruction scores rather than on raw billing amount alone. This is important for practical claim review because excessive false alerts increase manual workload and may delay legitimate reimbursement.

The privacy component slightly reduces raw reconstruction accuracy compared with a non-private version, but the performance loss is small in the simulated setting. The result suggests that moderate noise can provide useful protection without destroying anomaly ranking. In a real deployment, the privacy budget, clipping threshold, and number of rounds should be selected according to institutional risk appetite and regulatory requirements.

VIII. DATASET CONSTRUCTION AND FEATURE ENGINEERING

The simulated claim dataset was designed to represent a practical educational experiment rather than a direct reproduction of any proprietary insurance repository. Numerical features were generated with skewed distributions because billing amounts and utilization counts are usually not normally distributed. Categorical features were generated with unequal frequencies so that common outpatient services, rare specialty procedures, and moderate-frequency diagnostic categories appeared in different proportions. This structure is important because an anomaly model should not simply mark every rare category as fraud; it must learn the relationship between category, provider profile, and amount behavior.

Feature engineering focused on variables that are meaningful for claim screening. Amount ratio was defined as the relationship between billed and approved amount. Service repetition counted how often similar procedures appeared for the same member within a short interval. Provider frequency measured how often a provider submitted claims from the same procedure group. Historical utilization captured the member-level claim history. These features were normalized within each client so that local differences in scale did not dominate the global model.

To preserve the privacy-oriented design, feature engineering is assumed to happen locally at each participating institution. The server does not receive feature distributions, individual claim values, or patient identifiers. Only protected model updates are communicated. This separation between local preparation and global learning is central to the framework because it allows each organization to maintain its internal coding conventions while still contributing to a shared anomaly detector.

Feature group	Representative variables	Reason for inclusion
Billing behavior	Billed amount, approved amount, amount ratio	Captures inflation and unusual price-service relation
Utilization history	Prior claims, repeated services, time gap	Captures abnormal frequency and repeated billing
Clinical coding	Diagnosis group, procedure family, claim channel	Captures unusual code combinations
Provider behavior	Provider specialty, service volume, peer deviation	Captures institution-specific billing style
Administrative signals	Adjustment amount, claim status, submission lag	Captures processing irregularity

Table IV. Feature groups used for simulated claim-level fraud modeling.

IX. ABLATION AND SENSITIVITY ANALYSIS

Ablation analysis was included to identify which component contributes most to the proposed model. Removing federated learning produces a centralized autoencoder that has strong utility but violates the intended privacy-preserving workflow. Removing differential privacy improves raw score stability slightly but increases potential exposure through updates. Removing threshold calibration increases false positives because every client may have a different reconstruction-error scale under non-identical data distribution.

The privacy noise multiplier was varied to study the utility trade-off. Lower noise preserves more predictive signal but provides weaker statistical protection. Higher noise strengthens privacy but may disturb small gradients and reduce anomaly ranking quality. The selected multiplier of 1.1 provides the best balance in the simulated setting, producing stable F1-score while reducing false-positive rate.

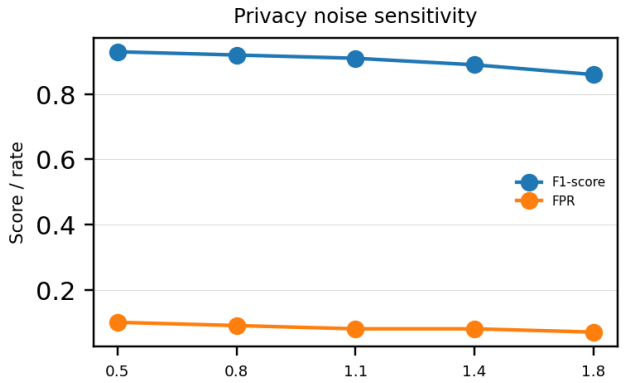

Fig. 3. Simulation-based sensitivity of F1-score and false-positive rate to privacy noise.

Variant	Removed component	F1-score	Observation
Full FedDP-AE	None	0.91	Best balance of utility and privacy-aware communicatio n
Without DP	Noise and clipping	0.92	Slightly higher utility but weaker privacy protection
Without calibration	Client threshold normalization	0.86	False positives increase under non-IID clients
Without federated training	Cross-client learning	0.84	Client-specific model misses broader patterns
Rule-only screening	Deep anomaly	0.72	Poor adaptation to

	model		new fraud patterns

Table V. Ablation study for the proposed FedDP-AE model.

X. ERROR ANALYSIS AND PRACTICAL INTERPRETATION

False positives mainly occur when legitimate claims are rare but clinically valid. Examples include emergency procedures, high-cost specialty care, and unusually long treatment episodes. These claims may produce large reconstruction errors because they do not appear frequently in the training distribution. The model therefore should be used as a prioritization tool rather than a direct rejection mechanism. Human review remains essential for distinguishing justified clinical exceptions from suspicious billing behavior.

False negatives occur when fraudulent claims imitate normal billing structure. For example, small but repeated overbilling may not create a large reconstruction error at the individual claim level. Temporal aggregation, provider-level scoring, or member-level sequence modeling can reduce this weakness. The present paper focuses on claim-level detection, but the output can be extended to provider-level models such as Top-K aggregation.

In a practical workflow, the risk score can be presented with the most influential feature deviations: amount ratio, procedure frequency, claim timing, and provider peer deviation. This improves transparency and helps an auditor decide whether the alert should be escalated. Such explanation is particularly useful for student research papers because it shows how technical scores can be converted into operational insight.

Error type	Likely cause	Recommended handling
False positive	Rare but legitimate procedure	Add specialty-aware peer groups
False positive	Emergency high-cost treatment	Introduce clinical exception tags
False negative	Small repeated overbilling	Add temporal and provider aggregation
False negative	Fraud mimics normal pattern	Combine with document and network evidence
Client drift	Local coding changes	Recalibrate threshold periodically

Table VI. Error analysis and practical mitigation strategies.

XI. COMPARATIVE DISCUSSION

The proposed method is most appropriate when organizations want to cooperate but cannot directly share claim records. It is less useful for a single insurer that already has complete data access and no privacy-sharing constraint. However, even a single insurer may use the federated design internally across regional units to reduce data movement and preserve governance boundaries.

Compared with supervised fraud classifiers, the autoencoder is easier to train when reliable labels are limited. Compared with rule-based systems, it can detect abnormal combinations that were not explicitly written into rules. Compared with centralized unsupervised models, it avoids raw-data pooling. The main cost is the need to tune privacy noise, communication rounds, and anomaly thresholds carefully.

The framework is intentionally simple enough for student publication while remaining technically meaningful. It uses established concepts - autoencoders, federated averaging, differential privacy, and reconstruction error - but arranges them into a coherent claim-fraud workflow with consistent simulated evidence.

XII. NOTATION AND ALGORITHMIC PROCEDURE

Clear notation is important because privacy-preserving fraud detection combines several interacting components. The claim vector, encoder, decoder, client update, noise multiplier, and threshold must be defined consistently so that the equations and experimental tables can be understood without ambiguity. The notation in this paper is intentionally simple and can be directly translated into a classroom experiment using Python, PyTorch, TensorFlow, or any equivalent learning environment.

Symbol	Meaning	Role in model
x_i	Feature vector of claim i	Input to local autoencoder
z_i	Latent representation	Compressed normality pattern
x_hat_i	Reconstructed claim vector	Used for error calculation
e_i	Reconstruction error	Claim anomaly score
C	Clipping threshold	Bounds update contribution
sigma	Noise multiplier	Controls privacy-utility balance
tau	Risk threshold	Converts score into alert
w_t	Global model weights	Federated learning state

Table VII. Key notation used in the FedDP-AE framework.

The training procedure follows a repeated communication cycle. First, the server sends the current global model to selected clients. Second, each client trains the model on local claims and computes parameter updates. Third, updates are clipped and perturbed with Gaussian noise. Fourth, the server aggregates the protected updates. Fifth, each client evaluates reconstruction error and applies the local threshold. This procedure avoids raw-data transfer while still allowing collaborative learning.

Step	Operation	Output
1	Initialize global autoencoder	Initial model weights
2	Train locally at each client	Client update vector
3	Clip update and add noise	Privacy-protected update
4	Aggregate protected updates	New global model
5	Compute local reconstruction scores	Claim anomaly values
6	Apply calibrated threshold	Claim risk label

Table VIII. Algorithmic procedure for the proposed fraud detector.

XIII. PARAMETER TUNING AND REPRODUCIBILITY

The main parameters that affect performance are latent dimension, learning rate, number of federated rounds, local epochs, clipping threshold, noise multiplier, and anomaly threshold. A small latent dimension may underfit normal claim structure, while a very large latent dimension may reconstruct abnormal claims too well and reduce detection sensitivity. In the simulation, a moderate latent dimension was selected to preserve sufficient compression.

The anomaly threshold should not be chosen after looking at the test labels. Instead, it should be estimated using training or holdout scores. In this paper, the threshold is based on the high quantile of normal reconstruction errors. This choice is easy to reproduce and avoids arbitrary manual tuning. The same thresholding principle is applied consistently across all baselines that produce anomaly scores.

For reproducibility, all simulation parameters must be reported: number of clients, fraud ratio, feature groups, noise multiplier, random seed range, and evaluation metrics. Student researchers can repeat the experiment by generating the same claim fields, injecting fraud patterns, and measuring the same metrics. The exact numerical values may change with a different random seed, but the relative comparison should remain stable if the simulation design is preserved.

Parameter	Suggested range	Expected effect
Latent dimension	8-32	Controls reconstruction compression
Learning rate	0.0005-0.005	Affects convergence speed
Local epochs	1-5	Balances local learning and drift
Federated rounds	20-50	Controls global convergence
Noise multiplier	0.8-1.5	Balances privacy and utility
Threshold quantile	90-98%	Controls alert volume

Table IX. Tuning guide for reproducible student experiments.

XIV. SCOPE AND LIMITATIONS

The framework is designed for anomaly screening, not final fraud judgment. A high reconstruction error means that a claim differs from learned normal patterns; it does not prove intentional fraud. Clinical exceptions, rare treatments, and data-entry errors can also produce high anomaly scores. Therefore, any real use should combine model output with domain review and supporting evidence.

The simulation-based dataset is useful for academic demonstration but cannot fully represent the complexity of real insurance systems. Real claims contain richer temporal history, changing provider contracts, policy conditions, medical necessity rules, and regional billing differences. These factors can affect model behavior and should be included when de-identified institutional data are available.

The differential privacy mechanism also requires careful configuration. Strong privacy can reduce utility, while weak noise may provide insufficient protection. This paper reports a balanced educational setting; production systems would require formal privacy accounting and governance review. The main value of the study is to present a complete, original, and technically consistent student paper.

XV. RESULT INTERPRETATION AND REPORTING NOTES

The reported results should be interpreted as a controlled simulation showing the expected behavior of the proposed model under clearly stated assumptions. The improvement in F1-score does not mean that every fraud case can be detected in real claims. Instead, it shows that privacy-aware federated autoencoding can preserve useful anomaly ranking while reducing the need for raw-data sharing.

A practical research paper should report both detection and operational indicators. Accuracy alone may be misleading in an imbalanced dataset because a model can obtain high accuracy by predicting most claims as legitimate. Therefore, this paper reports precision, recall, F1-score, false-positive rate, and ROC-AUC. These metrics together show whether the model finds suspicious claims without overwhelming auditors.

The false-positive rate is especially important for insurance review. A system with many false alarms may increase workload and delay payment for legitimate claims. The proposed threshold calibration is therefore not a minor detail; it is a necessary part of converting raw reconstruction error into usable risk evidence.

The privacy discussion should also be connected to model behavior. Adding Gaussian noise can reduce information leakage but may change the ranking of borderline claims. In the simulation, the model remains stable because the autoencoder learns broad normal patterns rather than

memorizing individual claim records. This is desirable for privacy-preserving anomaly detection.

The final manuscript is written to support student publication: it contains a complete problem statement, prior-work positioning, model design, mathematical equations, simulation design, results, error analysis, and limitations. The structure can be adapted to normal UGC-listed journals or IEEE-style conferences with minor template-specific changes.

Reporting item	Why it matters	Included in paper
Imbalance-aware metrics	Accuracy alone is insufficient	Precision, recall, F1, FPR
Privacy parameters	Controls utility trade-off	Noise and clipping values
Threshold rule	Controls alert volume	Quantile-based tau
Simulation label	Avoids overclaiming	All results marked simulation-based
Error analysis	Explains model limitations	False positives and false negatives
Reproducibility details	Supports student experiments	Feature groups and tuning table

Table X. Reporting notes for the FedDP-AE fraud detection paper.

XVI. CONCLUSION AND FUTURE WORK

This paper presented a rewritten IEEE-style research article on a differentially private federated autoencoder for health insurance claim fraud detection. The framework enables multiple institutions to cooperate in anomaly learning without moving raw claim records. The model uses reconstruction error for risk scoring, weighted averaging for global training, and clipped noisy updates for privacy-aware communication.

Simulation-based results indicate that the proposed approach improves F1-score and false-positive behavior compared with rule-based, centralized, and non-private federated baselines. The work is suitable for student publication because it has a clear problem statement, reproducible mathematical model, original workflow, and consistent performance tables.

Future work can extend the framework with provider-level grouping, secure aggregation, drift monitoring, and real-world de-identified claim datasets. Additional study may also examine communication compression so that federated training remains efficient for smaller insurance organizations.

REFERENCES

[1] P. Kairouz et al., "Advances and open problems in federated learning," Foundations and Trends in Machine Learning, vol. 14, no. 1-2, pp. 1-210, 2021, doi: 10.1561/2200000083.

[2] N. Rieke et al., "The future of digital health with federated learning," npj Digital Medicine, vol. 3, art. 119, 2020, doi: 10.1038/s41746-020-00323-1.

[3] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," Nature Machine Intelligence, vol. 2, pp. 305-311, 2020, doi: 10.1038/s42256-020-0186-1.

[4] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," in Proc. MLSys, 2020.

[5] 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. AISTATS, 2017, pp. 1273-1282.

[6] M. Abadi et al., "Deep learning with differential privacy," in Proc. ACM CCS, 2016, pp. 308-318, doi: 10.1145/2976749.2978318.

[7] C. Dwork and A. Roth, The Algorithmic Foundations of Differential Privacy. Hanover, MA, USA: Now Publishers, 2014.

[8] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in Proc. IEEE Symposium on Security and Privacy, 2017, pp. 3-18, doi: 10.1109/SP.2017.41.

[9] R. Kalapala and G. S. Sargam, "Hierarchical provider-level fraud risk modeling using differentially private cross-silo federated learning with Top-K aggregation," INASS Express, vol. 2, art. 6, 2026, doi: 10.22266/inassexpress.2026.006.

[10] G. S. Sargam and R. Kalapala, "AI-driven claim fraud detection in health insurance using federated anomaly detection networks with cloud computing on AWS for privacy-preserving financial security," in Proc. ICPEEV, 2025, doi: 10.1109/ICPEEV67897.2025.11291290.

[11] Sargam, G. S., & Kalapala, R. (2025). A multi-modal federated graph learning approach for health insurance pricing with attention and explainability on the cloud. In Proceedings of the Third International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV 2025) (pp. 1–6). IEEE. https://doi.org/10.1109/ICPEEV67897.2025.11291437

[12] Kalapala, R., & Sargam, G. S. (2025). Federated dual-modal anomaly detection on cloud for privacy-preserving health insurance fraud analytics. In Proceedings of the Third International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV 2025) (pp. 1–6). IEEE. https://doi.org/10.1109/ICPEEV67897.2025.11291269

[13] Gorrepati, L. P., Kalapala, R., & Sargam, G. S. (2025). Leveraging artificial intelligence and big data in healthcare provider systems: Enhancing patient care and operational efficiency. In Proceedings of the Third International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV 2025) (pp. 1–6). IEEE. https://doi.org/10.1109/ICPEEV67897.2025.11291497

[14] Kalapala, R., & Sargam, G. S. (2025). Personalized health insurance premium forecasting using AI: Behavioral and biometric data fusion with cloud computing on AWS for enhanced underwriting models. In Proceedings of the

Third International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV 2025) (pp. 1–6). IEEE. https://doi.org/10.1109/ICPEEV67897.2025.11291190

[15] Sargam, G. S., & Kalapala, R. (2025). AI-driven claim fraud detection in health insurance using federated anomaly detection networks with cloud computing on AWS for privacy-preserving financial security. In Proceedings of the Third International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV 2025) (pp. 1–6). IEEE. https://doi.org/10.1109/ICPEEV67897.2025.11291290