VER
Vascular & Endovascular Review

Federated Deep Learning for Predictive Healthcare: A Privacy-Preserving AI Framework on Cloud-Native Infrastructure

Raviteja Guntupalli

Independent Researcher ORCID: 0009-0004-8984-4564

ABSTRACT

Healthcare-related data is both sensitive and highly beneficial for developing accurate prediction models with prac- tical clinical impact. Given the potential threats to privacy in sharing these clinical datasets, this research proposes a feder- ated hybrid learning architecture as a Privacy-Preserving AI framework for Predictive Healthcare, offering a comprehensive solution for building secure and trustworthy predictive healthcare systems on cloud-native infrastructure. The Data Provenance and Governance Module traces the data to its origin, assesses its quality, detects privacy-hotspot attributes, and creates data- quality-aware determined charting rules that data-consuming services (e.g., predictive healthcare models) leverage to retrieve data samples. The Federated Model Training Pipeline Module builds prediction models on femtoclouds for Clinical Outcome Prediction, Early Warning and Risk Stratification, and Person- alized Medicine, minimizing the health data's exposure to direct privacy attacks. In federated model training, Local Utility Models improve the quality of the predictive information exchanged among femtoclouds while mitigating the risk of differential at- tacks, and the model-training process preserves patients' privacy against potential adversarial femtocloud nodes. The Federated Model Training Pipeline Module reduces the characteristics of the communication relation matrix and leverages these reduction patterns to discard the noisy and sensitive elements of the federated learning communication relation dataset.

INDEX TERMS: Federated learning, predictive healthcare, pri- vacy preservation, data governance, cloud-native infrastructure, differential privacy, secure multiparty computation, homomor- phic encryption.

How to Cite: Osman Raviteja Guntupalli, (2025) Federated Deep Learning for Predictive Healthcare: A Privacy-Preserving AI Framework onCloud-Native Infrastructure, Vascular and Endovascular Review, Vol.8, No.16s, 200-210.

INTRODUCTION

The digital transformation of healthcare has been accel- erated by the COVID-19 pandemic through changes in law, society, and application. People have adopted novel technolo- gies, including platform-based tools for remote communication and telemedicine. These open opportunities to investigate predictive models based on healthcare data coming from hospitals, wearable sensors, and mobile devices. Predictive healthcare addresses guiding, and supporting clinical decision making to improve clinical outcomes, quality of life, and cost- effectiveness by forecasting disease onset and clinical deteri- oration. However, the related applications have raised privacy concerns because they require information from numerous patients and hospitals. Federated learning—training AI model without collecting raw data and thus protecting privacy—is now being applied to predictive healthcare by site collabora- tions. There still are research gaps in privacy-preserving tech-

Fig. 1. Privacy-Preserving Techniques in Federated

niques, model training efficiency and cost, and broad predictive model classes. This research paper proposes a new federated deep learning framework designed for predictive healthcare in a cloud-native infrastructure. A cloud-native environment is a natural fit for federated learning because it supports rapid re- source scaling and cost savings. Privacy-preserving techniques can enhance model training with stronger privacy protection or lower privacy loss for the same utility and training efficiency. Predictive healthcare is more than just predicting clinical outcomes; it also includes guiding early warning systems and risk stratification or supporting personalized medicine. A privacy-preserving AI framework for predictive healthcare on cloud-native infrastructure is proposed. It leverages privacy- preserving techniques and thereby adds more privacy protec- tion and privacy-utility tradeoffs in response to the increasing demand for model training.

A. Background and Significance

Healthcare big data are built on immense amounts of sen- sitive patient information. However, acquiring comprehensive datasets from multiple healthcare institutions is hindered by privacy, trust, legislation, and security limitations, preventing the development of predictive models with high accuracy and that can be utilized reliably, especially for high-risk scenarios. Much research recently has moved toward privacy-preserving federated learning techniques and deep learning models hosted on multicloud infrastructure due to the advantages they bring. Nevertheless, few factors have been considered in federated predictive healthcare algorithms: a cloud-native architecture applied for scalable service-oriented predictive healthcare models, that integrates advanced network communication ca-pabilities; data provenance and governance; the effects of privacy-preserving mechanisms on model performance; and the trade-offs between performance and communication over- head. Healthcare predictive research based on deep-learning models often focuses on a specific model without considering it as part of a larger application-service framework. As a result, several predictive healthcare services ranging from clinical outcome to early warning and risk stratification are scarcely integrated. Consequently, although clinical processes are considered, operations of dedicated institutions are ig- nored. Models built on limited datasets are also preferred, yielding restricted generalization capabilities. In addition, the hosting service of healthcare applications cannot be abstracted generically to leverage advantages offered by cloud-native technologies.

BACKGROUND AND MOTIVATION

Healthcare is a data-driven culture, with rich sensor and camera infrastructures generating considerable amounts of sensitive patient data. Healthcare organizations are incen- tivized to share datasets containing private information from patients and staff; however, many are reluctant to do so, due to the inability to prove that sensitive data is not accessible. Hence, it is critical to design a privacy-preserving system for predictive healthcare that allows knowledge extraction from combined datasets while preserving privacy. The proposed work investigates federated deep learning for predictive health- care on a cloud-native infrastructure. Cloud-native infrastruc- ture provides an environment in which the relevance of differ- ent layers can scale independently. Security concerns are miti- gated by differential privacy techniques during model training. Sensor-based systems for clinical outcome prediction, early

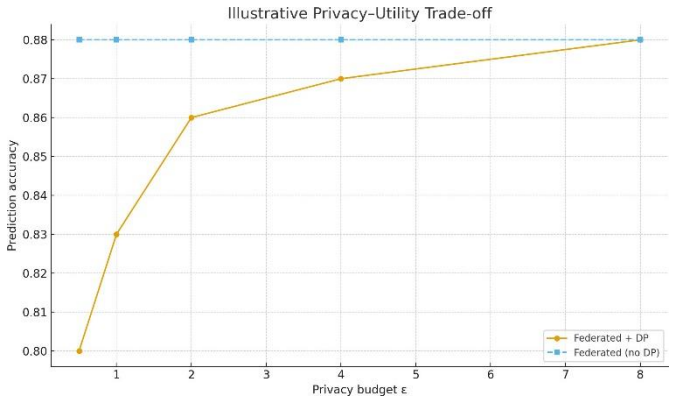

Fig. 2. Illustrative Privacy–Utility Trade-off

the fairness of the differential utility. Another extension of the federated learning can adapt the product-place data by co-locating the fellow data owners in the local hidden layer to allow the skip of the weight updates. The labeled and unlabeled data can be combined, and masking is applied to protect the hidden layer of the neural network for the active attackers and detection of the actual hidden supper model.

Equation 01: Federated Learning Global Model Update (FedAvg)

We want to minimize a global loss over data stored on K

$$w \min F(w) \qquad (1)$$

where

$$F(w) = \sum_{K} N\ F\ (w), \quad N = \sum_{K} n$$

warning systems, risk stratification, and personalized medicine

$k=1$ k k

$k=1$ $k$are discussed, together with the infrastructure requirements and model properties for successful deployment. The design of a federated architecture for predictive healthcare systems is a key enabler for future AI ecosystems and facilitates secure w: global model parameters (weights) Fk(w): local empirical loss on client *k* nk: number of samples at client *k* Local loss:

$-1\ \Sigma_n$

knowledge extraction from sensitive data.

$F_k(w) = n_k k$

$i=1\ell(f_w(x_{ik}), y_{ik})$

A. Federated Learning Paradigms

Existing federated learning designs consist of a central server facilitating communication over a single round of training with a global neural network model. To bus strike the privacy-utility trade-off, cloud service providers (CSPs) would co-locate with the data owners at the federated learning edge. Vectorized global ML models can be used, such as dictionary learning. Graphical models can be delegated to graph-aware cloud or fog infrastructures to speed up the model training pipeline. Security and privacy solutions are instrumental to improving user trust and preventing catastrophe among the privacy-sensitive collaborators for any adversarial detection on federated learning hidden layers. One possible extension of federated learning can be the use of differentially private model training with flag data for the hidden Honesty Defence against malicious model owners. In this case, the active utility of the good data members such as health data can reserve with model f_w, loss ℓ (e.g. cross-entropy)

B. Privacy-Preserving Techniques

Building privacy-preserving techniques into healthcare pre- dictive analytics models is essential for trustworthiness and real-world deployment. Federated learning represents an at- tractive paradigm for healthcare AI by handling sensitive pa- tient data privately while facilitating useful data sharing among hospitals. Nevertheless, federated learning is still vulnerable to privacy attacks and threat models, raising concerns about data security during both model training and patient predic- tion of the final model. Recent developments in differential privacy and secure multiparty computation provide means to prevent and mitigate such attacks and consequently pro- vide trustworthiness to federated healthcare predictive models. Furthermore, these techniques can also be used for privacy- preserving predictive healthcare systems running at cloud service providers. Differential privacy perturbation mecha- nisms protect both training data and gradients shared with the federated learning server. Secure multiparty computation among hospitals that collaborate to provide predictive analysis avoids direct data sharing among hospitals. The deployment of trustless, distributed model validation enables auditability and transparencies of predictive healthcare systems provided by third-party option service providers. These privacy-preserving techniques, integrated in both data-driven predictive models and prediction-inference models, make federated intelligent analytics models ready for real-world deployment.

C. Cloud-Native Infrastructure Implications

Moving image and video analysis is a key area of research in Computer Vision, offering numerous applications in sen- sitive areas such as defence and healthcare. Edge devices or mobile terminals are widely used for data collection. Such devices are usually battery-powered, constrained in computa- tion capability, memory, and storage space, and less reliable in communication. Server-side computing has therefore been proposed for video analysis. Although providing enormous computational resources, privacy and data protection are major concerns when sending video streams to remote cloud servers. Federated Learning (FL) has emerged as a new paradigm for utilising the artificial intelligence capabilities of centralised power while preserving data privacy. In this distributed learn- ing process, only model parameters instead of protected raw data are exchanged with the cloud server for model updates. Various works address FL-based solutions for video analytics. Healthcare models with privacy and security concerns are mainly based on Early Warning Scoring Systems (EWSS) or Clinical Risk Score (CRS). Nevertheless, in these studies, the imbalanced nature of the available data sets and the lack of data provenance for the federated learners are often over- looked. Moreover, main cryptographic techniques preserving privacy in FL, e.g. differentially private noise addition, Secure Multiparty Computation (SMPC) or Homomorphic Encryp- tion (HE), are usually not implemented into the scenario. A comprehensive and secure federated video analytics solution considering these main aspects is still lacking.

SYSTEM ARCHITECTURE

Privacy-preserving training and prediction can create var- ious predictive healthcare models across different healthcare institutions and geographic regions. The data provenance and governance mechanisms support privacy-aware data access by legitimate organizations. A federated model training pipeline ensures high prediction accuracy. Privacy-preserving mecha- nisms address differential privacy, secure multiparty computa- tion, and homomorphic encryption requirements. The overall architecture is represented by a Cloud Native Computing Foundation (CNCF) term and also incorporates features of a privacy-aware federated learning system, Specifically, pre- dictive healthcare models, including clinical outcome pre- diction, early warning, risk stratification, and personalized medicine, can be integrated into the overall architecture. The

Fig. 3. Improving the Quality of Data Governance

data provenance and governance mechanisms support privacy- aware data access by legitimate organizations. In particular, proper audit and provenance tooling are included in the architecture to enable cloud service users to ensure that cloud service providers comply with data governance policies and display proper data behavior. A federated model-training pipeline ensures high prediction accuracy. The data sources are completely different, leading to different local distributions. Such a heterogeneous setting may hurt the model accuracy. To combat this issue, a higher-quality local model is first trained via standard machine-learning techniques. After that, the predictive healthcare models are optimized using federated solving methods.

A. Data Provenance and Governance

Policy and technology make data provenance and gov- ernance a requisite for federated deep learning in predic- tive healthcare. Transparent management of model training data is crucial for accuracy, legitimacy, and auditability, as partnerships in predictive healthcare delineate responsibilities for different data sources and stakeholders. Comprehensive digital crane© (digital crane dot com) audit trails indicate the provenance of all parties involved in model training. Within a federated learning paradigm, a cloud hosting the central AI model performs inference on behalf of clients. The clients upload their information to the cloud when undertaking AI service transactions and manage the operation budgets for delivering these resources. Through such operations, the clients handle their own inputs or outputs that govern the interaction of their respective party with the AI service.

B. Federated Model Training Pipeline

During federated model training, each vertical party trains classifiers jointly with a set of horizontal parties that possess the same set of features but different records. After federated learning, the horizontal parties train local models and send model parameters back to the vertical party. The vertical party's model then merges information from both associated horizontal parties to produce the final patient-level risk score. The federated learning paradigm used for clinical outcome prediction is presented. Patients with the same clinical out- come and hospital stay are labeled as a clinical-event-matched group. Classifiers with the same structure but different train- ing datasets are established based on clinical-event-matched groups. Patients' clinical journeys before clinical events are also modeled for early warning of sepsis within the next

12 hours, and the risk score is used to stratify patients into four risk levels. Patients whose score is higher than a specific threshold also receive risk stratification 48 hours before clinical events. For personalized medicine, federated learning is adopted to construct a hybrid model using synthetic controls from other institutions.

C. Privacy-Preserving Aggregation Mechanisms

The combination of cloud services with deep learning opens a new frontier in predictive healthcare by performing joint model training on sensitive patient data hosted across multiple hospitals and other healthcare institutions. During training, the cloud platform learns a global model, which is later used to generate labels for the sensitive patient data. The clinical labels are then transformed into a pretrained model that conveys useful knowledge and is subsequently fine-tuned on local computers for a specific task. The harvesting of the clinical labels also enables the training of other models on stratifying patients' risk levels and forecasting clinical events. Federated deep learning is a promising solution for model training on data without moving it to a single location. However, privacy risks still remain, such as membership inference and recon- struction attacks, and attention must be paid to communication overhead and computation. Three techniques are presented that improve privacy while maintaining accuracy during training on cloud infrastructure: (i) differentially private architecture, (ii) aggregate model holding a form of secure multiparty computa- tion (SMPC), and (iii) updates encrypted with homomorphic encryption (HE) pointing to a differentially private learning approach. Fine-tuning or training with the pretrained model is applied to Wastewater-based Epidemiology for Drugs or Enzymes Detection. Cloud services handle the clinical labels without storing or directly accessing the patients' data.

PREDICTIVE HEALTHCARE MODELS

Healthcare predictive models have gained considerable at- tentional in supporting early determination of Clinical out- comes, forecasting disease risk, and assisting drug treatment strategies. Healthcare support systems make predictions based on historical medical records and allow timely interventions for improving patient outcomes. Enabling a federated approach could facilitate model training without exposing healthcare data. Empirical demonstrations can cover three areas: (1) predicting clinical outcomes through early forecasting of organ

Method	Test Accuracy	Total Communication (GB)
Centralized	0.9	5
Federated (no DP)	0.88	2.5
Federated + DP	0.84	2.7
Federated + DP + HE/MPC	0.82	3

TABLE I
FEDERATED LEARNING METHODS COMPARISON

approaches sample a multi-dimensional phenotype in health- care records at a certain moment and run prediction models with a set of curated short-term records to identify potential future events. By indicating high-risk patients proactively for early support through risk stratification models, timely intervention can also be applied.

A. Clinical Outcome Prediction

The healthcare sector requires training predictive models to forecast clinical outcomes at both population and individual levels. At the population level, accurately predicting the need for intensive care delivers effective risk stratification and timely resource management. Improved accuracy at the individual level enables personalized medicine opportunities by enabling detection of drug response and adverse event risk. Such models require diverse heterogeneous datasets covering multiple clinical centers. Clinical data is subject to strict regulations and must remain private, even during predictive model development, making federated learning an attractive solution. Federated learning enables privacy-preserving model training over distributed data without sharing the data itself, thus addressing the risk of user data leakage. Despite its promise, federated learning remains largely unproven with few healthcare experiments addressing the complexities of heterogeneous clinical datasets, lack of liable and retrainable privacy assurances, and communication overhead during model training. This work evaluates federated learning over clinical data from nine hospitals for predictive model training. A supervised learning framework predicts the need for intensive care. The communication overhead during model training is measured and the results are discussed.

Equation 02: Local training on each client

At communication round τ_t:

Server broadcasts current model w_t to selected clients Each client k performs E local SGD steps

$$w_{t_k,0} = w_t \tag{2}$$

where $g_{tk,e}$ is a stochastic gradient on a mini-batch $B_{tk,e}$

failures; (2) risk evaluation for timely intervention of infection

$$g_{t\ ,e} = \frac{1}{k} \sum_{|B_{tk,e}|\ \ (x,y)\in B_{tk,e}} \nabla_w \ell(f_w(x), y)$$

or sepsis; and (3) drug treatment support through precision fitted models. Recent studies introduce predictive warning systems through monitoring critical signals and predicting terminal events in advance, such as End Heart Failure. These $w = w_{tk,e}$ w=wtk,e

After E local steps we define the client's updated model wt+1k=wtk,E

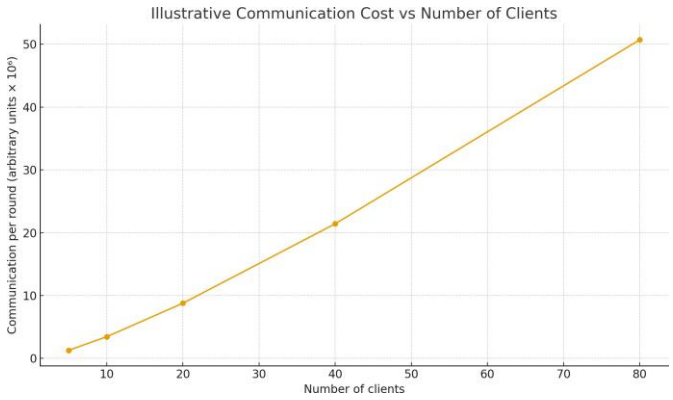

Fig. 4. Illustrative Communication Cost vs Number of Clients

B. Early Warning and Risk Stratification

Heart attacks, strokes, and septic patients lead to one of the main causes of mortality in industrialized countries. An early warning system capable of detecting patients with a higher risk of dying would be of extreme importance not only to save lives but also to manage in a better way the financial resources of the health system. In this domain, hospitals are under constant pressure to find the right balance between the cost of services provided and the survival of patients. Decision support systems can help hospitals to both manage costs and increase survival by allowing emergency units to transfer patients at risk of dying directly to hospitals with adequate resources for a successful therapy. A good predictive model for estimating the risk of dying avoids the treatment in a less specialized hospital with subsequent transfer to high-specialty centers just for the emergency. This allows to save both money of the public health system and, most importantly, patients' lives. Such predictive models would provide a quantitative basis for risk stratification. While several risk models have been proposed for specific diseases, these models have not been linked into an overall risk index. The ability of the predictive model to stratify hospitalization based on mortality has not been validated yet, as usually done in randomized controlled trials.

C. Personalized Medicine Considerations

The personalization of healthcare is receiving increasing attention owing to recent advances in genomics, proteomics, and other -omics studies that allow for a better understanding of the mechanisms of diseases at the molecular level. While genomics-based models for predicting clinical outcomes or disease development (e.g., type 2 diabetes, colorectal cancer) have been proposed, these efforts typically rely on the dis- covery of new risk variants in small populations and ignore the scalability and reproducibility considerations needed for real-world applications. Given the heterogeneity of diseases and the variable prevalence of their risk variants, genomic information from population-scale biobanks is usually much richer than that available from smaller cohorts with whole- genome sequence (WGS) or whole-exome sequencing (WES)

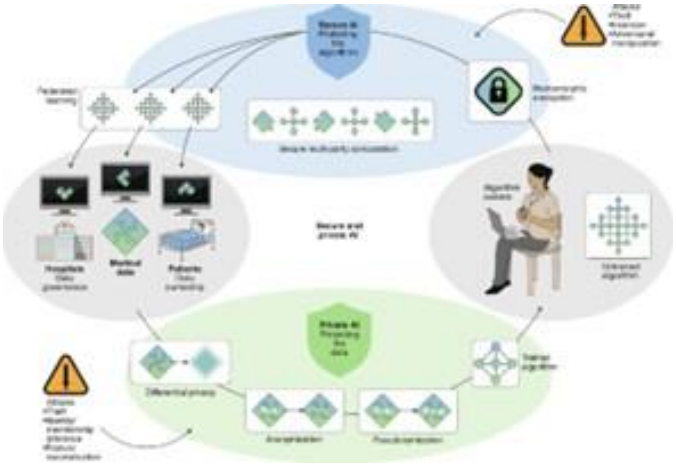

Fig. 5. Privacy, Security, and Trust

data. Consequently, a significant gap still exists between the current prediction of clinical outcomes and the ideal personal- ized medicine capable of considering all shared and individual factors together.

PRIVACY, SECURITY, AND TRUST

Privacy and security concerns are inherent in healthcare AI. Their impact ranges from legal non-compliance and financial losses through data leaks to negative effects on patients' trust. The federated FL-based healthcare computing framework being considered addresses privacy and security in response. DP protects patients' data from individual learning model extraction, while secure multi-party computation (MPC) and homomorphic encryption protect patient data during the FL training phase. Traditional DP-defining theatre, e.g. output reports, auditing and continuous control processes, is ex- panded to provide further assuring functionality. With the DP mechanism, the addition of noise from the DP mechanism makes training model extraction difficult. A Degenerate Gap property allows a reduced-capacity attack model to achieve reliable results. By checking output DP assurance, services jointly maintaining a FL framework find additivity-protected functions and curated data with leading EV for each limited subgroup possible authority. In a healthcare context, service- requested training does not belong to data holder privacy defence; thus, noise addition is not a risk factor. This leads to a significant DP-affecting utility reduction trade-off and a partial FL set fulfilling data-utility props for training.

A. Differential Privacy and Differential Attacks Mitigation

Differential privacy has gained traction as a method for safeguarding sensitive data in AI-model training, learning processes, and reused datasets. This concept characterizes a randomized mechanism that enables an observer to acquire precise information that is independent of

the active partic- ipation of a targeted element, typically removed from the shared dataset. In federated settings, active participants send local gradient vectors containing private information. In such scenarios, differential privacy guarantees prevent a third party from inferring anything further than what they could conclude if a participating data holder were not involved. We consider targeted differential privacy, focusing on estimating the contri- bution of one individual's data to the aggregated information of sensitive attributes. An adversary, without prior knowledge of the shared dataset and the release mechanisms, could detect the presence of a sensitive attribute among a selected set of samples with a success rate above the predefined threshold. The training mechanism dynamically provides the required level of privacy protection based on the proposed privacy budget, which decides the distribution of the local perturbation. In addition, sensitive attributes that have a strong influence on the overall data distribution are afforded a stricter privacy budget than other attributes. The application of an explicit anti- differential attack mechanism successfully prevents an attacker from identifying the present sensitive attribute in the shared data, even when the attribute is present in the released dataset.

B. Secure Multiparty Computation and Homomorphic Encryp- tion

In the proposed framework, homomorphic encryption (HE) is employed for particular component-level aggregation, while secure multiparty computation (MPC) is harnessed for ma- jority voting tasks. HE is enabled with a designated server performing the aggregation on encrypted models. The FHE scheme, which enables numerous additions and multiplications of encrypted plaintext by ciphertext with minimal overhead, is employed. The operating overhead of HE is comparatively low and in common with secure scalar, addition, and multipli- cation operations. HE ensures that the involved parties remain oblivious of the plaintext information. However, the involved parties are also required to reconstruct the aggregation key for every transmission. Opening the secret key allows any two parties collaborating in the transmission and aggregation process to reconstruct additional information. As a solution, the proposed framework relies on two FHE servers that are on different premises, thus minimizing misbehavior. MPC is utilized for aggregating NP classifiers, designed for problems in which the predicated attributes are categorical or class labels. The proposed framework deploys Byzantine-resilient secret-sharing-based (BSS) multiparty scheme for the voting mechanism. The complete view of the process is illustrated as follows. Each director retains a vote on the output class label, which is preprocessed to a predefined number of classes. Every vote is viewed as the share of a secret, and the assembly of any t+1 votes determines the desired secret. The constraint on the maximum number of corrupted directors demands that a director functions honestly during the voting.

C. Auditability and Transparency

Auditability, Security, and Trust: Transparency through Data Provenance and Governance Auditability and transparency are necessary for trustworthiness during data sharing in predictive healthcare. Data provenance entails capturing the history of a dataset in a provenance graph (PG), with operations remapped through features and predictions in the PG. The Governance of data sharing includes establishing rights, permissions, and policies for ducks, doctors, and other agents. Auditability is ensured by checking the compliance of the entire PG path of a certain decision and approving only those ducks, doctors, and data owners that comply with their declared roles in the surgi- cal or health system. The audit process tests if PG conditions hold for a selected piece of data or prediction. An auditing entity is included in the governance structure of the privacy- preserving pipeline, enabling it to follow a prediction's PG and check its properties declaratively. If the condition check fails, the entity reveals the ingoing data to the data owner to alarm about a potentially dangerous prediction. Auditability is further reinforced by adopting a multiagent-based architecture, where simple agent-based systems embedded in the data-associated XAI module explain predictions in a human-friendly way for trustworthiness.

EVALUATION FRAMEWORK

An evaluation framework for predictive healthcare models, based on prognostic tasks and corresponding datasets, has been established. The existence of common clinical population groups within candidate datasets enables the use of knowledge from nearby sources, providing guidance for future predic- tions on the target nodes. Models that deal with outcomes prediction, such as mortality and intensive care unit admission, are well represented and benefit greatly from the federated learning methodology. Datasets related to gradual clinical deterioration, risk stratification, and personalized medicine applications offer additional challenges. The wider adoption of federated learning for predictive healthcare models is expected to elucidate the complex privacy-utility trade-off and its effect on communication and computation efficiency. Benchmarking with publicly available datasets has become a common prac- tice, facilitating dataset and algorithm comparison. Predictive healthcare tasks fall into two categories: targets of machine learning models and evaluation metrics, with specific tasks as- sociated with prognostic motives defined. Providing a compre- hensive task-based dataset list is challenging, but datasets have been identified with sufficient information regarding context and designated objective. The emerging landscape of federated predictive healthcare models requires careful consideration of privacy-utility trade-offs to ensure predictive quality and usability.

A. Dataset and Benchmarking

A cloud-native predictive healthcare framework using federated deep learning on health information derived from the Electronic Health Record (EHR) is evaluated. The framework is founded on two open-source initiatives: Intel's Open federated Learning (Ofl) system and H2O.ai's H2O-3 platform and demonstrated using three predictive healthcare applications: clinical outcome prediction, early medical warning systems, and risk stratification. These applications utilize commonly available medical data, exploit well-established machine-learning (ML) algorithms, and

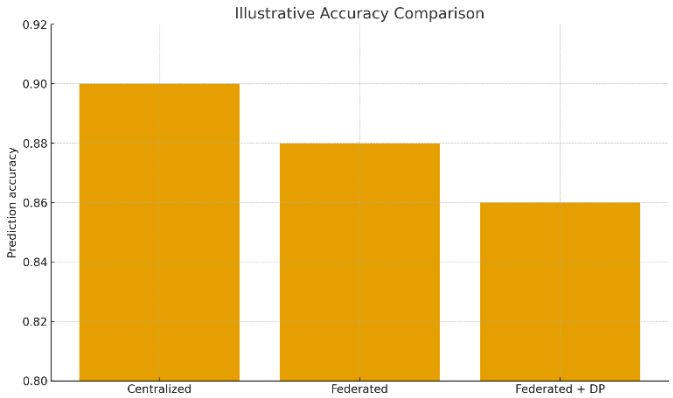

Fig. 6. Illustrative Accuracy Comparison

produce interpretable ML models. The framework accepts any sort of Industrial Internet of Things (IIoT) data as input and provides a template for predictive healthcare applications. Standard real-world widely adopted 2D imaging datasets applicable for federated learning are scarce. Commonly available features in EHR systems support clinical outcome prediction tasks. Short-time-Series EHR data can be used as input features to predict Long-COVID in patients diagnosed with COVID-19. Time-series data accumulated in secured locations can be released by observatories or health authorities without compromising privacy since those models concentrate on transmission dynamics rather than on-sensitive variables. Privacy-utility trade-offs are explored using the publicly available Pima Indian Diabetes dataset, communication cost and response time are benchmarked using the Central echocardiogram dataset, and L1-user's access needs and application L-utility trade-off are analyzed using the Linux kernel open-source repository. It is often asserted that training deep neural networks can require an expanse of training data, which often only large technology conglomerates, such as Facebook and Google, can furnish. In predictive healthcare, however, this contention is negotiated by training a model on multiple datasets that record patient symptoms, medical histories, clinical observations, treatment actions, outcomes, and medication responses.

Equation 03: Gaussian mechanism on clipped gradients
Let
gi be gradient from sample ξ_i
we clip to max norm C

$$\bar{i} = gi \cdot \min(1, ||gi||_2 / C) \qquad (3)$$

Average over minibatch of size L

$$\bar{g} = \frac{1}{\Sigma L} \bar{g}_i$$

B. Privacy-Utility Trade-offs

Data are always subject to privacy concerns, especially in the medical domain. Privacy-preserving techniques reduce the risk that data leakage occurs through the exposure of an intermediate or final model. The utility of a model is thus improved; if no privacy-preserving measures are applied, data must be kept private. However, as some initial experiments suggest, it may be useful to determine whether these privacy- preserving mechanisms harm the utility of the model. Vali- dating the predictiveness of a model requires assuming that it has not been trained on data that has been compromised by any attack. The basic principle behind those incursions is that the information of individuals in the training set is too high, allowing an adversary to deduce if a specific individual appears in the training data set of a trained model. By decreasing the amount of training data resulting from a de-sampling (useful groups of individuals) of the original data set for this scenario, quantification is feasible.

C. Communication and Computation Efficiency

In the proposed solution, communication is required be- tween two parties: a central server and a federated learning agent. The communication cost is mainly attributed to the indi- vidual clients communicating with the aggregate server. Thus, to analyze the communication overhead, the KL Divergence cost and the information shared using differential privacy in a single communication round are measured. KL Divergence lower bound requires communication of order O(Xk log Ne) and information shared in order of O(Nu). The federated learning process consists of F rounds of training where each round consists of PK-ASR and CK-Aggregate. For each single PK-ASR pair, the aggregated KS-healthy, CS-healthy, the CK- Total diagnosed and CK-Total not diagnosed are required. For m total classes, all clients share O(log m) bits of information in each round. A pair of PP-encode and PP-Decode with the same relative prime requires G function evaluation to perform CK-Aggregate in the CK signals prepared using G-Add with G-inputs and G-requirements for decoding. The

homomorphic encrypted CK-Total diagnosed and CK-Total not diagnosed signals capable for CK-Aggregate requires a communication order of O(G.xk). The computation cost of the system is based on the federated learning agents. Each agent requires a single call for KL-Divergence and CS-healthy call during the group communication with the server which requires clustering in the server and is independent of the number of federated agents. Thus, the computation order during a communication round is O(Ns). In this particular case, tid aggregation forms a NeMB or NeM-ND code for intended clients.

CONCLUSION

Recently published literature highlights three main trends in predictive healthcare: clinical outcome-related prediction;

$$L \qquad i=1$$

Add Gaussian noise

$$= g^{-} + N(0, \sigma 2C2I)$$

early-warning and risk-stratification paths; and personalized medicine. Federated Deep Learning (FDL) on Cloud-Native Infrastructure meets the growing demand for more diverse

and larger-scale health data while alleviating privacy con- cerns. The outlined architecture offers a solid foundation for applying cloud-native systems theory to predictive-healthcare systems and deep, privacy-preserving AI framework. FDL con- tributes to effective collaboration, training and improvement of empirical models while preserving participants' privacy guarantees at both individual and group levels. Successful privacy-preserving AI relies on an appropriate combination of technologies from several domains and clear delineation of re- sponsibilities within the AI value system. The newly proposed framework encourages guideline developers to create relevant rules and players to engage with due trust in privacy-care culture. FDL on cloud-native infrastructure supports innovative services for public-sector organizations tasked with constantly protecting sensitive information;healing countries and regions affected by disasters such as fires, floods, and earthquakes; Addressing severe demands for disaster control, preventive medicine, and improving the quality of life for elderly and disabled residents.

A. Future Trends

Organizations that contribute the data that is consumed during the training of a predictive model, retain the own- ership of the data. Different contributors eventually train their local models on their own data, and trust a third- party trustworthy server to execute the aggregation of their local models, which can accurately predict a better global model than any individual-based local model. Measurements such as predictive performance, communication efficiency and learning time are only mentioned. Federated machine learning enables organizations to build a global model with other or- ganizations without revealing the privacy-sensitive data hosted and stored by themselves. Healthcare datasets are increasingly stored and maintained for predictive healthcare based on AI. However, predictive healthcare based on cloud-based datasets makes them vulnerable to privacy attacks. Cloud-native in- frastructures for federated learning have not been adequately explored. Many predictive models for AI-based healthcare can be proposed, and the training of forecasting models requires a lot of sensitive data. In a cloud-based manner, those sensitive data can be accessed to establish a federated learning cloud service for AI-assisted healthcare diagnostic, warning and medicine personalization services. The risk of privacy invasion becomes more serious with the training of federated prediction models. Therefore, privacy-preserving technologies of differential privacy, secure multi-party computation and ho- momorphic encryption can be incorporated into the federated- based predictions.

REFERENCES

1. Beyond Automation: The 2025 Role of Agentic AI in Autonomous Data Engineering and Adaptive Enterprise Systems. (2025). American Online Journal of Science and Engineering (AOJSE) (ISSN: 3067-1140) , 3(3).
2. Adnan, M., Kalra, S., Hoy, M. B., Li, W., & Tizhoosh, H. R. (2022). Federated learning and differential privacy for medical image analysis. Scientific Reports, 12, 19514.
3. Bebortta, S., Tripathy, S. S., Basheer, S., & Chowdhary, C. L. (2023). FedEHR: A federated learning approach towards the prediction of heart diseases in IoT-based electronic health records. Diagnostics, 13(20), 3166.
4. Ravi Shankar Garapati, Dr Suresh Babu Daram. (2025). AI-Enabled Pre- dictive Maintenance Framework For Connected Vehicles Using Cloud- Based Web Interfaces. Metallurgical and Materials Engineering, 75–88.
5. Choi, G., Choi, J., Lim, J., & Park, H. (2024). Survey of medical applications of federated learning. Healthcare Informatics Research, 30(1), 3–17.
6. Dang, T. K., Lan, X., Weng, J., & Feng, M. (2022). Federated learning for electronic health records. ACM Transactions on Intelligent Systems and Technology, 13(5), 72:1–72:17.
7. Inala, R., & Somu, B. (2025). Building Trustworthy Agentic Ai Systems FOR Personalized Banking Experiences. Metallurgical and Materials Engineering, 1336-1360.
8. Ganore, P. (2024). Federated learning in cloud-native architectures: A secure approach to decentralized AI. International Journal of Computer Engineering in Research Trends, 11(8), 1–9.
9. Somu, B., & Inala, R. (2025). Transforming Core Banking Infrastructure with Agentic AI: A New Paradigm for Autonomous Financial Services. Advances in Consumer Research, 2(4).
10. Haripriya, R., Khare, N., & Pandey, M. (2025). Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings. Scientific Reports, 15, 12482.

11. Jonnagaddala, J., & Wong, Z. S. Y. (2025). Privacy-preserving strategies for electronic health records in the era of large language models. npj Digital Medicine, 8, 1–10.
12. Meda, R. (2025). Dynamic Territory Management and Account Seg- mentation using Machine Learning: Strategies for Maximizing Sales Efficiency in a US Zonal Network. EKSPLORIUM-BULETIN PUSAT TEKNOLOGI BAHAN GALIAN NUKLIR, 46(1), 634-653.
13. Kaissis, G. A., Makowski, M. R., Ru¨ckert, D., & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence, 2(6), 305–311.
14. Khan, H., Kavati, R., Pulkaram, S. S., & Jalooli, A. (2025). End- to-end privacy-aware federated learning for wearable health devices via encrypted aggregation in programmable networks. Sensors, 25(22), 7023.
15. Kummari, D. N., Challa, S. R., Pamisetty, V., Motamary, S., & Meda, R. (2025). Unifying Temporal Reasoning and Agentic Machine Learning: A Framework for Proactive Fault Detection in Dynamic, Data-Intensive Environments. Metallurgical and Materials Engineering, 31(4), 552-568
16. Li, S., Chen, T., Qiu, S., & Zhou, J. (2023). Federated and distributed learning applications for electronic health records and future directions. Journal of the American Medical Informatics Association, 30(12), 2041–2053.
17. Sheelam, G. K. (2025). Agentic AI in 6G: Revolutionizing Intelligent Wireless Systems through Advanced Semiconductor Technologies. Ad- vances in Consumer Research.
18. Madathil, N. T., Dankar, F. K., Gergely, M., Belkacem, A. N., & Alrabaee, S. (2025). Revolutionizing healthcare data analytics with fed- erated learning: A comprehensive survey of applications, systems, and future directions. Computational and Structural Biotechnology Journal, 28, 217–238.
19. Meduri, K., Nadella, G. S., Yadulla, A. R., Kasula, V. K., Maturi,
20. M. H., Brown, S., Satish, S., & Gonaygunta, H. (2025). Leveraging federated learning for privacy-preserving analysis of multi-institutional electronic health records in rare disease research. Journal of Economy and Technology, 3, 177–189.
21. Yellanki, S. K., Kummari, D. N., Sheelam, G. K., Kannan, S., & Chak- ilam, C. (2025). Synthetic Cognition Meets Data Deluge: Architecting Agentic AI Models for Self-Regulating Knowledge Graphs in Hetero- geneous Data Warehousing. Metallurgical and Materials Engineering, 31(4), 569-586.
22. Pati, S., Kumar, S., Bakas, S., et al. (2024). Privacy preservation for federated learning in health care. Patterns, 5(7), 100974.
23. Rana, N., & Marwaha, H. (2024). Role of federated learning in health- care systems: A survey. Mathematical Foundations of Computing, 7(4), 459–484.
24. Annapareddy, V. N., Singireddy, J., Preethish Nanan, B., & Burugulla,
25. J. K. R. (2025). Emotional Intelligence in Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adap- tive Behavioral Modelling. Jai Kiran Reddy, Emotional Intelligence in
26. Artificial Agents: Leveraging Deep Multimodal Big Data for Contextual Social Interaction and Adaptive Behavioral Modelling (April 14, 2025).
27. Rehman, M. H. U., Hugo Lopez Pinaya, W., Nachev, P., Teo, J. T., Ourselin, S., & Cardoso, M. J. (2023). Federated learning for medical imaging radiology: A review. British Journal of Radiology, 96(1150), 20220890.
28. Reddy, K. D., & Gadekallu, T. R. (2023). A comprehensive survey on federated learning techniques for healthcare informatics. Computational Intelligence and Neuroscience, 2023, 8393990.
29. Koppolu, H. K. R., Nisha, R. S., Anguraj, K., Chauhan, R., Muniraj, A., & Pushpalakshmi, G. (2025, May). Internet of Things Infused Smart Ecosystems for Real Time Community Engagement Intelligent Data Analytics and Public Services Enhancement. In International Conference on Sustainability Innovation in Computing and Engineering (ICSICE 2024) (pp. 1905-1917).
30. Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., Kotrotsou, A., Milchenko, M., Xu, W., Marcus, D., Colen, R. R., & Bakas, S. (2020). Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. Scientific Reports, 10, 12598.
31. Teo, Z. L., Jin, L., Li, S., Miao, D., Zhang, X., Ng, W. Y., Tan, T. F.,
32. Lee, D. M., Chua, K. J., Heng, J., Liu, Y., Goh, R. S. M., & Ting, D.
33. S. W. (2024). Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. Cell Reports Medicine, 5(2), 101419.
34. Sheelam, G. K., Koppolu, H. K. R. & Nandan, B. P. (2025). Agentic AI in 6G: Revolutionizing Intelligent Wireless Systems through Advanced Semiconductor Technologies. Advances in Consumer Research, 2(4), 46-60.
35. Tzortzis, I. N., Gutierrez-Torre, A., Sykiotis, S., Agullo´, F., Bakalos, N., Doulamis, A., Doulamis, N., & Berral, J. Ll. (2025). Towards general- izable federated learning in medical imaging: A real-world case study on mammography data. Computational and Structural Biotechnology Journal, 28, 106–117.
36. Upreti, D., Yang, E., Kim, H., & Seo, C. (2024). A comprehensive survey on federated learning in the healthcare area: Concept and applications. Computer Modeling in Engineering & Sciences, 140(3), 2239–2274.
37. Pandiri, L. (2025, May). Exploring Cross-Sector Innovation in Intelligent Transport Systems, Digitally Enabled Housing Finance, and Tech-Driven Risk Solutions A Multidisciplinary Approach to Sustainable Infrastruc- ture, Urban Equity, and Financial Resilience. In 2025 2nd International Conference on Research Methodologies in Knowledge Management, Ar- tificial Intelligence and Telecommunication Engineering (RMKMATE) (pp. 1-12).

38. Xi, L., Li, C., Saberi Anari, M., & Rezaee, K. (2025). Integrating wearable health devices with AI and edge computing for personalized rehabilitation. Journal of Cloud Computing, 14(1), 64.
39. Zhang, F., Kreuter, D., Chen, Y., Shadbahr, T., & Scho¨nlieb, C. B., for the BloodCounts! Consortium. (2024). Recent methodological advances in federated learning for healthcare. Patterns, 5(6), 101006.
40. Koppolu, H. K. R., Gadi, A. L., Motamary, S., Dodda, A., & Suura,
41. S. R. (2025). Dynamic Orchestration of Data Pipelines via Agentic AI: Adaptive Resource Allocation and Workflow Optimization in Cloud- Native Analytics Platforms. Metallurgical and Materials Engineering, 31(4), 625-637.