

Federated Transfer Learning Models for Privacy-Preserving Predictive Diagnostics in Heterogeneous Healthcare Systems Using Distributed Cloud Computing Frameworks

Gerald de Basto D,

Software Consultant, USA.

Citation: Gerald de Basto D. (2025). Federated transfer learning models for privacy-preserving predictive diagnostics in heterogeneous healthcare systems using distributed cloud computing frameworks. *International Journal of Advanced Research in Cyber Security (IJARC)*, 6(3), 1-6.

Abstract

The growing need for predictive diagnostics in healthcare demands both high-performing models and strict privacy assurances. This study presents a federated transfer learning (FTL) approach deployed within a distributed cloud computing framework to facilitate diagnostic predictions without centralized data collection. Addressing the heterogeneity of healthcare data across institutions, our method ensures privacy-preserving collaboration while enhancing model generalizability. We evaluate this approach on synthetic and real-world datasets (MIMIC-III and COVIDx), demonstrating its efficacy in maintaining diagnostic accuracy and minimizing data exposure.

Keywords: Federated learning, transfer learning, privacy preservation, healthcare diagnostics, cloud computing, model generalization, heterogeneous systems.

1. Introduction

The healthcare sector is undergoing a digital transformation where predictive diagnostics play a central role in early detection and personalized treatment planning. However, the integration of machine learning (ML) models in clinical settings is hindered by strict data privacy regulations (e.g., HIPAA, GDPR), institutional data silos, and infrastructural disparities. Traditional centralized training mechanisms are increasingly unsuitable due to the sensitivity and heterogeneity of medical records.

To address these challenges, federated learning (FL) has emerged as a solution that enables collaborative model training without centralizing patient data. Yet, FL faces limitations in model generalizability when applied to diverse, non-IID (independent and identically distributed) healthcare data. Transfer learning, particularly cross-domain adaptation, can mitigate this by allowing models to transfer learned knowledge from data-rich institutions to data-scarce ones. When embedded in distributed cloud computing

environments, federated transfer learning (FTL) achieves scalable, privacy-aware diagnostics across heterogeneous systems.

This paper explores the design and deployment of FTL models in cloud-based federated systems to preserve privacy while improving diagnostic accuracy in real-world healthcare applications.



2. Literature Review

Several studies have advanced the field of federated and transfer learning in healthcare.

2.1. Li et al. (2020) introduced *FedAvg*, a foundational algorithm for federated learning, demonstrating its ability to perform decentralized model training while ensuring data privacy. However, it underperformed on heterogeneous data.

2.2. Sheller et al. (2020) implemented FL for brain tumor segmentation across institutions. Their results confirmed that federated models could match the performance of centrally trained models, but they highlighted challenges in non-IID data convergence.

2.3. Zhang et al. (2021) applied transfer learning to COVID-19 detection, showcasing improved diagnostic outcomes in underrepresented patient groups, especially when combining domain adaptation with convolutional neural networks (CNNs).

2.4. Xu et al. (2022) proposed a hybrid federated transfer learning framework for cardiology data. They used residual blocks and fine-tuning layers to enhance model robustness across different healthcare settings.

2.5. Al-Rakhmi and Al-Qurishi (2023) evaluated privacy-enhancing technologies (e.g., differential privacy, homomorphic encryption) in federated learning systems. Their work established a trade-off between privacy preservation and model accuracy.

These foundational studies demonstrate both the promise and challenges of integrating FL and TL in healthcare, particularly around model divergence, communication overhead, and privacy assurance.

3. Methodology

This study proposes an **FTL framework** leveraging a distributed cloud computing infrastructure. The model architecture is built using a base CNN with domain-adaptive layers, pre-trained on MIMIC-III and fine-tuned across several synthetic hospital datasets representing varied patient populations.

3.1. System Architecture:

The system uses a **star-topology cloud network**, where each hospital (client) trains a local model on-site and sends model updates (not raw data) to a central aggregator.

3.2. Model Workflow:

- **Stage 1: Pretraining** on a large labeled dataset (source domain).
- **Stage 2: Local Fine-tuning** at each institution using transfer learning.
- **Stage 3: Aggregation** using weighted FedAvg with transfer-aware gradients.

Table 1. Dataset Characteristics

Dataset	Size (records)	Domain	Label Type	Source Institution
MIMIC-III	58,976	ICU	Mortality Risk	PhysioNet
COVIDx	16,756	Radiology	COVID+/COVID-	Open Repository
Synthetic A	10,000	Cardiology	Readmission Risk	Simulated Hospital A
Synthetic B	12,000	Oncology	Survival Status	Simulated Hospital B

4. Experimental Design and Metrics

To evaluate the effectiveness of our FTL framework, we conducted simulations using both real and synthetic datasets across five federated nodes.

4.1. Metrics:

- **Accuracy, Precision, Recall, and F1-Score** for classification tasks.
- **Communication Efficiency:** Number of rounds to convergence.
- **Privacy Leakage Score** using model inversion attack simulations.

4.2. Experiment Setup:

- **Hardware:** Google Cloud TPUs and NVIDIA A100s.
- **Frameworks:** TensorFlow Federated, PySyft (for privacy simulation).
- **Privacy Mechanisms:** L2 norm clipping, differential privacy noise injection.

5. Results and Analysis

Our model showed strong performance under heterogeneous conditions. Compared to baseline federated models (FedAvg, FedProx), the FTL model achieved superior diagnostic accuracy and faster convergence.

Table 2. Model Performance Comparison (Averaged Across Nodes)

Model	Accuracy	Precision	Recall	F1-Score	Rounds to Converge
FedAvg	81.2%	78.9%	76.3%	77.6%	55
FedProx	83.1%	80.4%	79.2%	79.8%	50
FTL (ours)	87.5%	85.9%	84.1%	85.0%	38

Privacy leakage simulations revealed that differential privacy successfully minimized risk with less than 1% model inversion success rate under our FTL scheme.

6. Conclusion

This study demonstrates that federated transfer learning deployed over distributed cloud infrastructures offers a viable path toward privacy-preserving and high-accuracy predictive diagnostics in heterogeneous healthcare environments. By leveraging pretrained models and local fine-tuning, institutions can collaborate effectively without compromising patient data. Future work should explore cross-modal transfer (e.g., text to image), optimize communication protocols, and investigate real-time deployment under stricter latency constraints.

References

- [1] Li, Tian, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. "Federated Optimization in Heterogeneous Networks." *Proceedings of Machine Learning Research*, vol. 100, 2020, pp. 429–450.
- [2] Subramanyam, S.V. (2025). Cloud-based enterprise systems: Bridging scalability and security in healthcare and finance. *International Journal on Science and Technology (IJSAT)*, 16(1), 1–20.
- [3] Sheller, Micah J., G. Anthony Reina, Brandon Edwards, Jason Martin, and Spyridon Bakas. "Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation." *BrainLes 2018*, Springer, 2020, pp. 92–104.
- [4] Zhang, Kai, Xin Song, Yujin Huang, Hui Wu, and Yuan Li. "Cross-Domain Transfer Learning for COVID-19 Diagnosis Using Chest X-Ray Images." *Computers in Biology and Medicine*, vol. 132, 2021, p. 104306.
- [5] Subramanyam, S.V. (2025). Revolutionizing enterprise workflows: The role of declarative rules in business process systems. *International Journal of Information Technology and Management Information Systems (IJITMIS)*, 16(2), 341–365.
- [6] Xu, Jing, Zhen Wang, Shuang Liu, and Kun Lin. "Federated Transfer Learning for Personalized Heart Disease Prediction." *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, 2022, pp. 1668–1679.
- [7] Al-Rakhami, Mona, and Mohammed Al-Qurishi. "Privacy-Preserving Machine Learning in Federated Healthcare Environments: Opportunities and Challenges." *ACM Computing Surveys*, vol. 55, no. 2, 2023, pp. 1–39.
- [8] Subramanyam, S.V. (2021). Cloud computing and business process re-engineering in financial systems: The future of digital transformation. *International Journal of Information Technology and Management Information Systems (IJITMIS)*, 12(1), 126–143.
- [9] Yang, Qiang, Yang Liu, Tianjian Chen, and Yongxin Tong. "Federated Machine Learning: Concept and Applications." *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, 2019, pp. 1–19.
- [10] Rieke, Nicola, Jake Hancox, Wenqi Li, Fausto Milletari, Holger R. Roth, Saeed Albarqouni, and M. Jorge Cardoso. "The Future of Digital Health with Federated Learning." *NPJ Digital Medicine*, vol. 3, no. 1, 2020, pp. 1–7.
- [11]

- [12] Subramanyam, S.V. (2023). The intersection of cloud, AI, and IoT: A pre-2021 framework for healthcare business process transformation. *International Journal of Cloud Computing (IJCC)*, 1(1), 53–69.
- [13] Kairouz, Peter, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Ahmad Al-Nasr, and Sen Zhao. "Advances and Open Problems in Federated Learning." *Foundations and Trends in Machine Learning*, vol. 14, no. 1–2, 2021, pp. 1–210.
- [14] He, Zhen, Xingxing Zhang, Linxing Xing, and Xiaoxiang Xie. "Personalized Federated Learning for Intelligent Health Diagnosis." *IEEE Internet of Things Journal*, vol. 8, no. 8, 2021, pp. 4476–4486.
- [15] Subramanyam, S.V. (2024). Transforming financial systems through robotic process automation and AI: The future of smart finance. *International Journal of Artificial Intelligence Research and Development (IJAIRD)*, 2(1), 203–223.
- [16] Kaissis, Georgios A., Markus R. Makowski, Daniel Rückert, and Rickmer F. Braren. "Secure, Privacy-Preserving and Federated Machine Learning in Medical Imaging." *Nature Machine Intelligence*, vol. 2, no. 6, 2020, pp. 305–311.
- [17] Tan, Chuanqi, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. "A Survey on Deep Transfer Learning." *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, 2018, pp. 2704–2710.
- [18] Rajkomar, Alvin, Jeff Dean, and Isaac Kohane. "Machine Learning in Medicine." *New England Journal of Medicine*, vol. 380, no. 14, 2019, pp. 1347–1358.
- [19] Kumar, K. (2020). Enhancing interpretability and explainability in deep neural networks for artificial intelligence. *QIT Press - International Journal of Artificial Intelligence (QITP-IJAI)*, 1(1), 1-4.
- [20] Priyadarsini, A. (2024). Advancing the understanding of representation learning in artificial intelligence systems. *QIT Press - International Journal of Artificial Intelligence (QITP-IJAI)*, 5(2), 1-5.
- [21] Bansal, A. (2020). Predictive modeling and complex system analysis reimaged through deep learning-powered artificial intelligence. *QIT Press - International Journal of Artificial Intelligence and Deep Learning Research and Development*, 1(1), 1-4.
- [22] Navya, M. (2024). Deep learning as the foundation for advanced cognitive automation and human-machine collaboration in artificial intelligence. *QIT Press - International Journal of Artificial Intelligence and Deep Learning Research and Development*, 5(2), 1-4.