Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 4442–4455, DOI 10.12694/scpe.v25i5.2905

FEDERATED LEARNING FOR INTERNET OF MEDICAL HEALTHCARE: ISSUES AND CHALLENGES

NIKITA CHELANI[*], SHIVAM TRIPATHY[†], MALARAM KUMHAR[‡], JITENDRA BHATIA[§], VARUN SAXENA[¶], SUDEEP TANWAR[‖], AND ANAND NAYYAR[**]

Abstract. Federated Learning is a decentralized machine learning method that allows collaborative model training across several devices or institutions while maintaining the privacy and localization of data. Since the raw data is used locally, this collaborative method enables the development of a strong and precise global model without jeopardizing the privacy and security of sensitive data. The healthcare sector is an important one that focuses on preserving and enhancing people's health through medical services, diagnoses, treatments, and preventative measures. Efficient evaluation of Federated Learning in the Internet of Medical Things (IoMT) enables breakthroughs in medical image analysis, electronic health record analysis, personalized treatment planning, and drug development by enabling institutions to train models locally on sensitive patient information without sharing raw data. This paper presents the role of Federated Learning in healthcare and current trends in Federated Learning-based healthcare. A case study is presented on deep Federated Learning for privacy-preserving in healthcare. Finally, challenges and future research directions are discussed in the paper.

Key words: Federated Learning, Healthcare, Data Privacy, Machine Learning, Medical Image Analysis, Electronic Health Records, Data Security.

1. Introduction. A powerful machine learning-based health protection system utilizes the doctor's clinical judgment and the computer's massive processing power. Machine learning is crucial in healthcare, especially in areas such as computer-aided diagnosis, image annotation, image-guided medical help, image database retrieval, multimodal image fusion, and medical image segmentation, where inadequacies may be fatal. There are more options to create a system for patient recovery thanks to the improvement of health-related information. In the healthcare sector, machine learning has had limited social impact. Machine learning is the key to decreasing healthcare costs and encouraging improved patient-clinician communication. Multiple health-related uses of ML solutions include assisting physicians in locating multiple patient-specific drugs and therapies and assisting patients in deciding when and if to arrange follow-up visits [2]. Various ML algorithms have been used in healthcare environment. It vary in their implementation in terms of their methodology, the nature of input and output data [37]. The authors in [38] reviewed IoT frameworks and ML algorithms in healthcare sector, specifically on voice pathology. It also covers the impact of ML in various healthcare applications such as blood pressure and oxygen saturation monitoring using smartphones.ML technologies continue to go deeper into the healthcare environment, which also places higher demands on intelligent healthcare [36]. Figure 1.1 shows the various applications of ML in healthcare.

Federated Learning is an innovative approach for training a global machine learning model using data scattered across many data groups, removing the necessity for raw data exchange. Once a global model is shared

[*]Department of Computer Science and Engineering, Institute of Technology, Nirma University Ahmedabad, Gujarat, India. (22mced03@nirmauni.ac.in).
[†]Department of Information Technology, L. J. Institute of Engineering and Technology, Ahmedabad, India. (spt3009@gmail.com).
[‡]Department of Computer Science and Engineering, Institute of Technology, Nirma University Ahmedabad, Gujarat, India. (Corresponding author, malaram.kumhar@nirmauni.ac.in)
[§]Department of Computer Science and Engineering, Institute of Technology, Nirma University Ahmedabad, Gujarat, India. (Corresponding author, jitendra.bhatia@nirmauni.ac.in)
[¶]Department of Computer Engineering, Govt. Mahila Engineering College, Ajmer, India. (vps@gweca.ac.in).
[‖]Department of Computer Science and Engineering, Institute of Technology, Nirma University Ahmedabad, Gujarat, India. (sudeep.tanwar@nirmauni.ac.in).
[**]School of Computer Science, Duy Tan University, Da Nang, Vietnam. (anandnayyar@duytan.edu.vn).

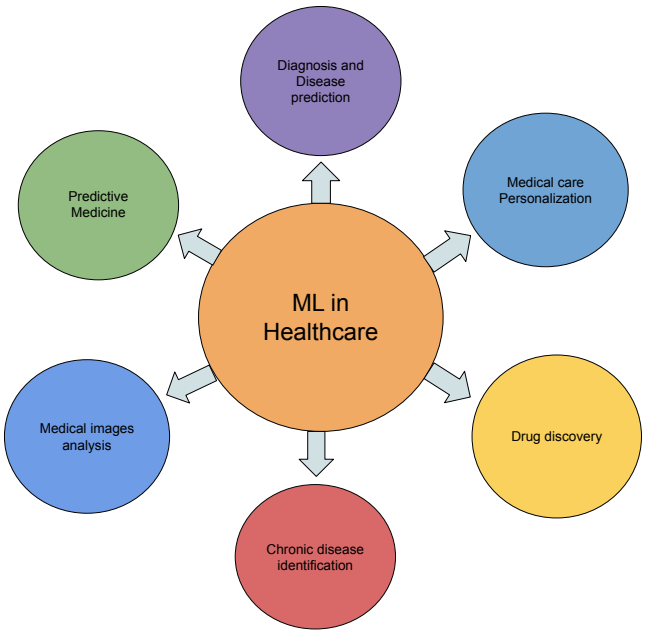

Fig. 1.1: Machine Learning applications in Healthcare

between different locations, the model is trained at each site using local data. The local models' parameter changes are then transmitted to an aggregation server and integrated into the final model. This is repeated until the global model's convergence requirement is met. Figure 1.2 presents the basic federated learning architecture. The benefits of Federated Learning have lately been proven in many practical applications, such as language modeling and picture classification. It is particularly pertinent in healthcare applications where data is replete with sensitive, personally identifiable information, and data analysis techniques must adhere to legal and regulatory standards [1].

1.1. Motivation. The potential for the Federated Learning technique to change the area of healthcare is what led to choosing it for research. The study aims to explore the intersection between cutting-edge machine learning methods and the complexity of healthcare data administration. The primary problem of protecting patient privacy and data security in healthcare research is addressed by Federated Learning's capability to support collaborative model training across diverse data sources. This compatibility between decentralized data analysis and the need for data secrecy provides a strong justification for this research work. The anticipated results, like improved diagnosis accuracy, customized treatment plans, and expedited medical research, highlight the profound change that Federated Learning might bring to the healthcare industry. Through the examination of Federated Learning applications, the research is conducted in hopes of assisting in the development of a healthcare environment that is more secure, effective, and patient-centered. Figure 1.3 shows various key technologies used for implementing Federated Learning in healthcare. Today's biggest difficulty for AI researchers and practitioners is how to legally address the issue of data fragmentation and isolation [3]. By developing a global model without distributing raw data among sites, Federated Learning offers a first degree of privacy protection. Federated Learning, however, may occasionally be exposed to inference attacks.

1.2. Scope of the paper. The importance of data privacy and security has emerged as a major global problem as huge organizations compromise user privacy and data security. Public media and governments are very concerned about reports of data leaks. Federated Learning aims to enable ML from non-colocated data, hence addressing privacy and data governance issues. Each data controller in an Federated Learning context establishes its governance procedures and privacy guidelines, managing data access and having the authority to withdraw it. It covers both the validation phase and the training phase. Federated Learning might provide new

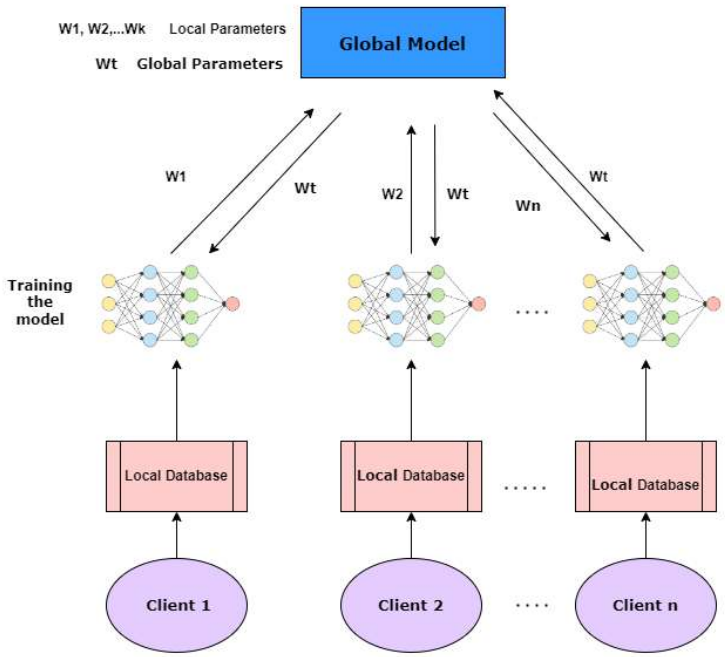

Fig. 1.2: Basic Federated Learning Architecture

opportunities in this way [4]. In this paper, the authors explore the role of Federated Learning in healthcare and review the current trends.

1.3. Organization of the paper. This section discusses the organization of the survey. The rest of the paper is organized as follows: Section 3 examines earlier research on collaborative machine learning that protects privacy with an emphasis on medical applications. Section 2 gives a brief introduction about the topic and its various use cases in medical field. Section 4 discusses the RPDFL method for digital healthcare applications.It deals with problems with conventional Federated Learning, such as gradient leaking and data silos in healthcare companies. Section 5 outlines machine learning's potential to benefit healthcare and its transformational role in clinical decision-making, healthcare research, and tailored therapy while maintaining data security and privacy.

2. Background. Technological breakthroughs have revolutionised the healthcare business over the last several decades, ushering in a new era of customised, efficient, and data-driven medical services. At the centre of this transformation are several interconnected technologies, including IoT, SDN, Fog Computing, IoMT, and Machine Learning. Each of these technologies is critical to the transformation of healthcare, the improvement of patient outcomes, and the simplification of administrative operations.

2.1. Healthcare. Healthcare is one of the industries that is connected to everyone's life. It is a necessary part of people's life, encompassing physical, emotional, and mental aspects. It entails detecting, treating, and dispensing vaccinations. The healthcare business has incorporated technological advancements to make it more efficient and effective. There are various stages in the evolution of healthcare that might be labelled as "Healthcare X.0." These stages represent significant developments and advancements in healthcare practices, techniques, and delivery systems. Figure 2.1 depicts the evolution of healthcare from 1.0 to 5.0, and the

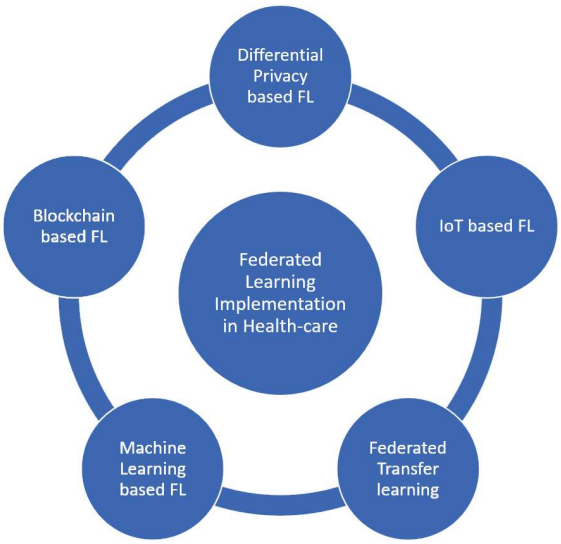

Fig. 1.3: Federated Learning Implementation in Healthcare

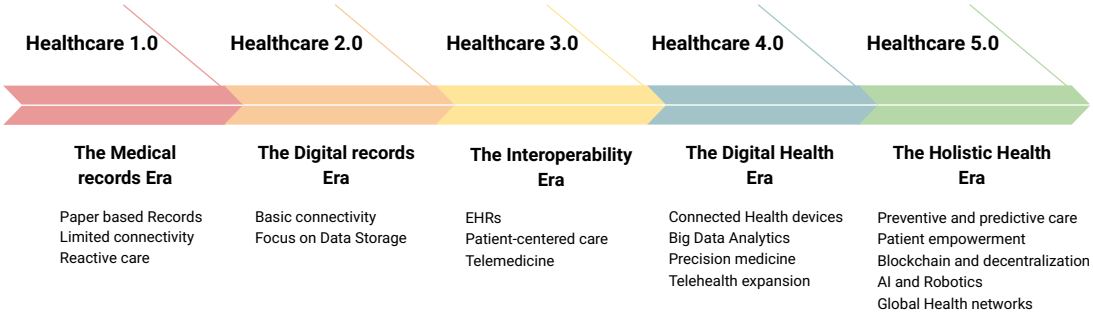

Fig. 2.1: Healthcare Domain Transformation [19]

significant advancements in healthcare are highlighted in this section.

- Healthcare 1.0: Healthcare was first focused on home medicines, traditional healing practices, and a lack of medical expertise. Medical treatments were not standardized, and healthcare was frequently given by local healers and herbal experts.
- Healthcare 2.0: This period saw improvements in medical knowledge, technology, and hospital development. During this time, the standardization of medical education and practices began. The doctor-patient connection was formalized.
- Healthcare 3.0: The digitization of healthcare records and the implementation of electronic health records (EHRs) occurred during this period. The emphasis has shifted to data-driven decision making and evidence-based medicine. Technologies for telemedicine and remote monitoring began to emerge.
- Healthcare 4.0: During this stage, big data, artificial intelligence, and machine learning are being integrated into medical practises. These technologies provide predictive and preventative treatment,

allowing for more personalised and accurate therapy. Because of the growth of wearable technology and sensors, patients may now measure their wellbeing and physical activity.

- Healthcare 5.0: Healthcare 5.0 is at the forefront of healthcare change. It focuses on personalized medicine, which involves adapting therapies to individuals' genetic, genomic, and lifestyle characteristics [21]. Advances in genomics, artificial intelligence, and big data analytics are critical in more precisely identifying and treating illnesses. Patient empowerment and active involvement in healthcare decision-making remain critical.

2.2. Federated Learning. Federated Learning, a revolutionary machine learning paradigm, has emerged as a possible answer to these problems in the healthcare industry. In contrast to traditional data aggregation methods, Federated Learning allows for the cooperative training of machine learning models across several institutions or devices while preserving the localization of raw data [24]. Data privacy is met by this dispersed strategy, which also protects patient data and complies with strict legal requirements like HIPAA. Federated Learning offers several benefits, including privacy preservation, reduced communication costs, and improved scalability [25].

- Privacy preservation: Since the data remains on the device, it is unnecessary to send it to a central server for processing, ensuring the data's privacy.
- Reduced communication costs: Since only model updates are sent to the central server, the communication costs are significantly lower than the traditional machine learning approach.
- Improved scalability: Since the computation is distributed among multiple devices, Federated Learning can handle multiple devices without needing extra resources.

The way people control their diabetes has been transformed by Continuous Glucose Monitoring (CGM) technology. For the treatment of diabetes, CGM systems continually assess glucose levels in real time. However, the volume and complexity of CGM data produced by these devices provide major hurdles, particularly in making accurate forecasts and enhancing patient outcomes. An original solution to these problems is the decentralized fusion of CGM data via Federated Learning, which combines the power of CGM data from many sources while maintaining individual privacy and security.

With the help of Federated Learning, CGM data owners may build a global glucose prediction model without disclosing their raw data, protecting their privacy [26]. Local updates from participating CGM devices are combined to train a global model. Diverse data sources are advantageous for this model aggregation. The global model can offer particular people specific insights and suggestions for managing diabetes. The decentralized method improves data security by lowering the possibility of data breaches or illegal access. The use of decentralized CGM data fusion and Federated Learning shows great potential for overcoming the difficulties in adequately maintaining and exploiting CGM data. It makes it possible to create precise, privacy-preserving models that provide people with diabetes the ability to take care of their health while advancing study and treatment in diabetes management. This technology has the potential to fundamentally alter how we perceive and manage diabetes as it develops. Figure 2.2 shows the process of implementing Federated Learning in healthcare.

2.3. Internet of Medical Things (IoMT). The Internet of Medical Things (IoMT) integrates medical devices with the Internet of Things (IoT). IoMT represents the future of modern healthcare systems, where all medical devices will be interconnected and monitored online by healthcare professionals [22]. This development promises to enhance the speed and reduce the costs of healthcare services as it progresses. IoMT enhances the volume of health data accessible to caregivers, diversifies its sources, and accelerates the processes of collection, transmission, and analysis. The increased flow of data aids in improving decision-making for both patients and healthcare providers. This network of interconnected technology allows for the constant collection, analysis, and transmission of health data, enabling real-time monitoring and management of patients' health. IoMT devices encompass wearable fitness trackers, smart implants, remote patient monitoring systems, and connected diagnostic instruments [20]. IoMT streamlines data flow and automates routine tasks, alleviating administrative burdens and enabling medical staff to concentrate more on patient care, thus boosting operational efficiency. It helps lower healthcare costs by decreasing hospital readmissions, reducing the need for in-person visits through remote monitoring, and optimizing resource allocation. IoMT also enhances personalized healthcare through comprehensive data, increases patient engagement with intuitive interfaces, and supports telemedicine, thereby

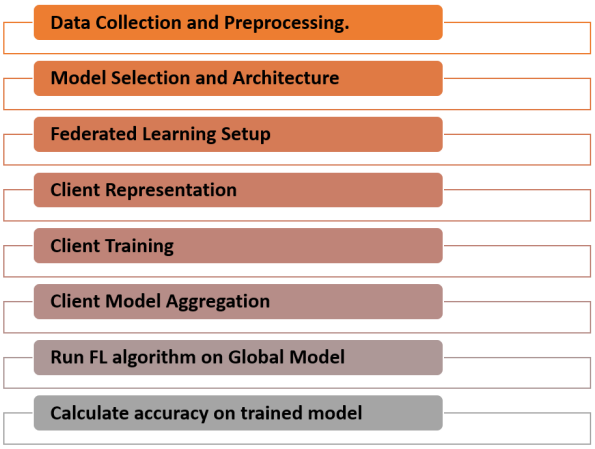

Fig. 2.2: General Implementation Process

improving healthcare accessibility, particularly in underserved regions [23]. Furthermore, IoMT ensures patient safety and adherence to prescribed treatments through reminders and alerts and enhances data analytics, propelling medical research and innovation.

3. Related Work. In the realm of Federated Learning, numerous studies have delved into the healthcare domain. This section reviews prior research on Federated Learning in healthcare, including new frameworks and architectures in this area and existing surveys.

Choudhury et al. [16] applied three classification algorithms—the perceptron, support vector machine (SVM), and logistic regression—that are amenable to distributed solutions using gradient descent for ADR and mortality prediction tasks. The models' usefulness is assessed using the F1 score before and after applying a privacy-preserving technique. It is demonstrated that FL, even without differential privacy, may deliver model performance near the theoretical scenario of total data centralization. The performance of the developed global FL model for a specific level of privacy is next examined with differential privacy. Contrasting the outcomes with baseline techniques demonstrates how the adaptive contribution weighting strategy enhances learning accuracy and convergence. The tests also validate the privacy-preserving techniques and show the framework's potential for usage in real-world smart healthcare applications.

Gupta et al. [8] examine using digital twins and hierarchical Federated Learning for anomaly detection in smart healthcare. It emphasizes the value of anomaly detection in healthcare and the difficulties with centralized, conventional methods. To enable data sharing and collaborative model training while preserving privacy, the authors suggest a hierarchical Federated Learning architecture that uses digital twins and virtual representations of actual items. The review gives a general description of the framework, goes through its benefits, and shows experimental findings to show how well it works at spotting abnormalities in healthcare data. The authors also examine possible uses for digital twins and hierarchical Federated Learning and potential future studies in smart healthcare.

Xu et al. [9] thoroughly investigate Federated Learning's use in healthcare informatics. It discusses Federated Learning's ideas, architectures, algorithms, and privacy-preserving techniques, among other topics. To illustrate the efficacy and promise of Federated Learning in various fields, the authors look at various healthcare informatics applications, including population health analysis, illness detection, therapy suggestion, and predictive modeling. The paper also emphasizes the significance of model aggregation approaches, data heterogeneity, and data quality in Federated Learning for healthcare. Overall, the literature analysis highlights research gaps, offers insightful information about the current status of Federated Learning in healthcare informatics, and makes recommendations for future developments.

Stripelis et al. [10] emphasize using Federated Learning to scale neuroscience research. Due to privacy issues and restricted access to massive datasets, they discuss the difficulties traditional centralized systems in neuroscience research confront. The paper demonstrates how Federated Learning might facilitate group analysis of dispersed neuroimaging data while protecting individual privacy. The authors also discuss the advantages of Federated Learning, including better data sharing, bigger sample sizes, and greater study reproducibility. Additionally, they give examples of Federated Learning applications in neuroscience, such as cognitive modeling, brain imaging processing, and research into neurodegenerative diseases.

Liu et al. [12] used a traditional method to assess the contributions of members inside a coalition accurately Shapley Value (SV). Furthermore, the coalition's overall utility is to be calculated because FL may use it as a fair contribution evaluation principle. The average of all marginal contributions a participant makes over-all coalitional permutations is that individual's SV. To investigate CAreFL's suitability for further intelligent healthcare application situations. Adding contribution-based data pricing mechanisms to the CAreFL framework would help the development of a market for FL-based healthcare data sharing. The ultimate objective is to include these capabilities in the free and open-source FATE framework so that they may be used by more programmers, scholars, and practitioners.

Le Sun and Jin Wu [13] proposed a model that comprises a parameter protection method to prevent catastrophic forgetting, a lightweight 1-D CNN-based feature extractor, and a PCC mechanism to facilitate class scaling. To increase the effectiveness of training models for new tasks, SCALT also implements two significant model transfer methods. The trained SCALT models' mean parameter values from previous tasks are utilized as model initialization for a new task. To eliminate fuzzy boundaries between various classes, a KL-divergence-based sharing knowledge removal technique is used in the transferring process.

Patel et al. [14] suggested FL in healthcare for improved privacy and security of user data. A centralized intelligence system must cope with many constraints, including resource constraints, data update delays, a lack of high precision and accuracy, and privacy and security issues. The suggested FL-EHR case study incorporates differential privacy, which guarantees learning data privacy, which is advantageous in real-world settings. It also eliminates data bias throughout the training phase, increasing the model's accuracy. As a result, the suggested study enhances model validation, the foundation for efficient analytics setups in diverse medical ecosystems compared to present healthcare systems. The FL training scheme calculates the local slopes. Then, using techniques like FedAvg, Fed-Prox, and FedSGD, all local gradients are combined at the aggregation level. Here, a 6G network is being used for communication. Effective optical wireless communication channels in 6G networks can be used to establish direct, close connectivity with healthcare infrastructure. The prediction model is then sent to numerous clients via the aggregation layer using a HE to FL training strategy.

Wu et al. [15] proposed FedHome, a cloud-edge FL system. FedHome employs a novel generative convolutional autoencoder architecture. To compensate for the loss in prediction performance caused by an unbalanced and non-IID distribution of user data, GCAE synthesizes minority class samples and retrains the user's local model using the resultant class-balanced dataset. The Federated Learning framework safeguards data privacy by keeping user data local and learning a common global model in the cloud from several network edge homes. To deal with the unbalanced and non-IID distribution inherent in user monitoring data, a generative convolutional autoencoder (GCAE) is created with the goal of achieving accurate and personalised health monitoring by refining the model with a generated class-balanced dataset from user-specific data.

Lu et al. [16] use a FedAP as a weighted, tailored, batch-normalized federated transfer learning technique for the healthcare industry. FedAP combines data from several companies while maintaining privacy and security, and by taking into account similarities and maintaining local batch normalization, it provides reasonably customized model learning. The statistical mean of the relevant layer's inputs and the running mean of the BN layer are positively correlated. As a result, the researchers may run many rounds of FedBN while maintaining local batch normalization and use the parameters of BN layers to substitute the statistics when a pre-trained model is unavailable. This variety is known as f-FedAP. The authors suggest FedAP, a personalized Federated Learning method for healthcare that uses adaptive batch normalization to collect data from several clients without sacrificing security and privacy and develop customized models for each client.

Islam et al. [17] gave a system for universal data sharing for Federated Learning that uses differentiated privacy to forecast specific heart failure conditions. The system minimizes the total data dimension through

feature selection, lowering the noise addition for differential privacy while keeping effective scores. The authors propose a feature selection strategy based on the correlation value of the data to reduce dimension and boost the utility of the ML models. The authors employ Federated Learning infrastructure to enable private data sharing among collaborating parties in a differentially private manner. The basic structure of this Federated Learning model is comprised of two components. A model manager with whom all data owners communicate and the owner(s) of the raw data used to train the model. Furthermore, the model is protected by a second privacy layer that employs differential privacy, and the authors assume that the aggregator server is truthful yet interested in recovering the raw data of a data owner from it. The same training data was utilized in both Naive Bayes and Random Forest in a two-party configuration where the data is divided into two groups. The data's correlation value serves as the foundation for feature selection. The authors then added noise using Laplace transformation to the statistics they had generated using Differential Privacy (DP) techniques. The generated noisy data are finally sent to the aggregator server in place of the raw data. The central aggregator server feeds the machine learning framework to create a global model that is used to forecast the likelihood of cardiac failure.

Li et al. [5] focus on developing a smart healthcare system that preserves privacy using Federated Learning techniques. The proposed approach aims to solve concerns about the privacy and security of sensitive patient data while utilizing the benefits of machine learning in healthcare. Patient data must frequently be centralized in traditional healthcare systems, raising privacy concerns and the possibility of data breaches. However, Federated Learning allows training models collaboratively without transferring raw data. Multiple edge devices, including wearables and smartphones, are included in the system to gather patient data.

Yang et al. [3] introduced the concept of federated machine learning (FML), and its different applications are explored. A distributed approach to machine learning called FML enables numerous parties to jointly train a single model without disclosing their personal information. The authors outline the FML systems' structure, consisting of a dispersed client base and a central coordination server. The study discusses methods and strategies to handle technological difficulties such as communication effectiveness, privacy protection, and system robustness. In addition, practical uses in healthcare, finance, and smart cities are investigated, demonstrating how FML promotes collaborative learning while protecting data privacy. Overall, the paper illustrates the importance of FML in dealing with privacy issues and thoroughly discusses its idea, design, difficulties, and practical applications.

Abdulrahman et al. [6] provides an in-depth overview of Federated Learning, tracing its development from centralized to distributed on-site learning and exploring its advancements. It covers various aspects of Federated Learning, including architectures, communication protocols, optimization algorithms, privacy-preserving mechanisms, and model aggregation methods. The survey discusses applications in healthcare, IoT, edge computing, and autonomous vehicles, highlighting challenges and open research areas. Additionally, it explores recent advancements such as federated transfer learning and reinforcement learning and emerging trends like edge intelligence and blockchain-based Federated Learning. The paper serves as a comprehensive resource for researchers and practitioners interested in understanding the current state and future directions of Federated Learning.

Prayitno and Shyu [7] thoroughly assess the literature on Federated Learning in the healthcare industry, emphasizing data characteristics and applications. To illustrate the usefulness and promise of Federated Learning in tackling healthcare issues, the authors look at various healthcare applications, including illness prediction, medical picture analysis, electronic health record analysis, and wearable device data analysis. The study intends to present insightful analyses of current research, identify knowledge gaps, and suggest future research avenues in Federated Learning in healthcare.

Table 3.1 summarizes the existing research work on Federated Learning in healthcare. This literature review on Federated Learning in healthcare finds an emerging area with a strong emphasis on protecting patient privacy and exploiting decentralized data for model training. To handle the variety of healthcare data, researchers are building specialized Federated Learning strategies focusing on multiple data types and matching machine learning models. Key use cases like illness prediction, medical picture analysis, and personalized therapy recommendations highlight the potential to improve patient outcomes and healthcare efficiency. However, issues such as communication overhead, model convergence, data quality, scalability, and regulation compliance

Table 3.1: Survey of Existing Research Works

Author	Year	Key Contributions	Evaluation Parameters	Limitations and Future Scope
Dai *et al.* [31]	2020	Presented a blockchain-enabled framework to address the security and privacy issues with IoMT systems	Security and Privacy	Improving the performance by implementing more scalable consensus algorithms is not easy. Deep learning can be used to enhance the performance
Choudhury *et al.* [11]	2020	Federated Learning framework with two levels of privacy protection that can learn a global model from dispersed health data.	The global Federated Learning model's performance is then evaluated with respect to differential privacy at a specific privacy level.	Differential privacy hampers Federated Learning's prediction capacity due to excessive noise.
Gupta *et al.* [8]	2021	A threat model for centralized anomaly detection with particular threat scenarios and suggested a privacy-preserving model.	Accuracy, sensitivity, and efficiency of the anomaly detection system.	Scalability, Data Quality and Algorithmic Challenges.
Xu *et al.* [9]	2021	Privacy-Preserving Techniques in Healthcare Informatics and Model Aggregation Approaches	Accuracy, privacy preservation, model performance, scalability, and Computational efficiency.	Real-world Deployments and Interoperability
Stripelis *et al.* [10]	2021	A brain age prediction model based on structural MRI scans from several locations with varying quantities of data and subject (age) distributions.	Elasped time and a Mean Absolute Error (MAE)	The proposed technique demonstrates distinct relative performance of the multiple training approaches and, in the future, federated transfer learning in neuroimaging might be investigated.
Ali *et al.* [34]	2022	Presented the role of FL in IoMT for detecting privacy threats	Privacy and Accuracy	Need robust and universal FL architecture from privacy perspective.
Arya *et al.* [35]	2022	Proposed an ensemble FL approach for training a DL model to classify decentralized data streams in IoMT	Accuracy, Precision, Recall and F1-score	Since ensemble learning is carried out on a server, so there can be a threat for data privacy.
Liu *et al.* [12]	2022	Adaptive contribution weighting mechanism, privacy-preserving methods, contribution-aware Federated Learning framework.	Accuracy and balance between performance and privacy.	CAreFL's potential for other healthcare applications
Le Sun and Jin Wu [13]	2022	A scalable and transferable classification framework, SCALT to protect privacy of petients data	Measured accuracy of ECG, EEG, and PPG data	SCALT's space complexity can be reduced by minimizing the feature extractor's volume. Sophisticated model transfer methods will enhance new model training.
Patel *et al.* [14]	2022	FL in HI for enhanced privacy and security and a centralized intelligence systems to overcome resource constraints, data delays, and privacy issues.	The FL-EHR case study incorporates differential privacy for data privacy and reduces bias during training, enhancing model accuracy and validation.	The framework examines FL performance on healthcare datasets with different sensitivity levels. It introduces a DP model for FL-HI, ensuring high secrecy and accurate diagnosis.
Wu *et al.* [15]	2022	Developed a Generative Convolutional Autoencoder (GCAE) to provide accurate and personalized health monitoring by improving the model using a newly constructed class-balanced dataset.	Local minibatch size, training passes on each client's data; and number of participating clients per communication round.	GCAE's few parameters reduce communication overhead during model transfer.
Lu *et al.* [16]	2022	FedAP is a weighted, batch-normalized federated transfer learning for healthcare.	BN layer parameters to replace statistics, enabling multiple rounds of FedBN with local batch normalization when a pre-trained model is not available.	Consider implementing more accurate methods for calculating and updating client similarity.
Islam *et al.* [17]	2022	Feature selection based on correlation values minimizes data dimension, preserving model effectiveness.	Model manager and data owner(s), with no direct message exchange between providers. Naive Bayes and Random Forest were trained.	Data anonymization and differential privacy should be included, and the score should be further enhanced using a better feature selection technique.
Li *et al.* [5]	2022	A convenient and privacy-preserving system named ADDETECTOR to detect Alzheimer's disease (AD)	Accuracy, Time overhead, F-score	Find additional useful features to reflect the characteristics of Alzheimer's disease and test the viability of ADDETECTOR on a bigger dataset.
Tian *et al.* [18]	2023	Ring Structure for Federated Learning, Threshold Secret Sharing Mechanism and edge-dropout handling	Accuracy and convergence pace	To boost parallelism to lower the computational cost and communication overhead of the RPDFL training scheme.
Rani *et al.* [32]	2023	An exhaustive survey on the security of FL-based IoMT applications	Data Security and Privacy	Real world health data analytics with integrity of the data.
Alamleh *et al.* [33]	2023	Developed an multicriteria decision-making (MCDM) framework for standardising and benchmarking the ML-based IDSs for FL architecture of IoMT applications	Systematic ranking and Computational cost	More investigation on MCDM methods that consider the experts' importance is needed.

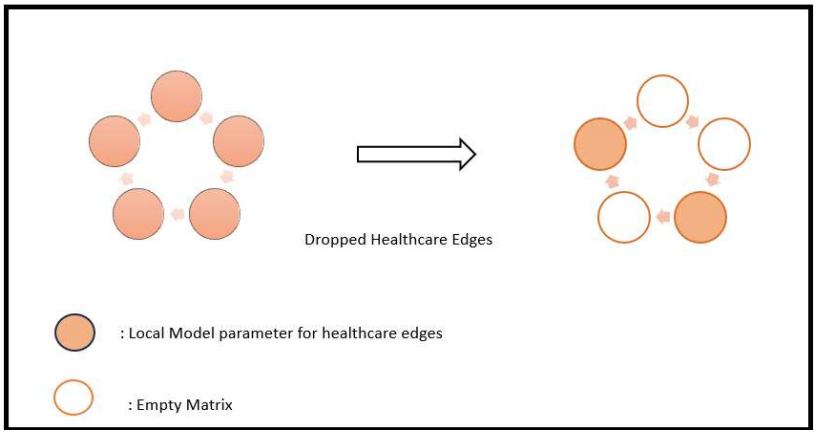

Fig. 4.1: Data sharing procedure of Ring AllReduce

prompt continuous research. Interdisciplinary research, the creation of Federated Learning frameworks, and an increasing interest in scalable and efficient methodologies are all exciting futures for Federated Learning's revolutionary role in healthcare.

4. Case Study. The potential to improve patient care and medical research has been demonstrated by using deep neural networks in digital healthcare applications. Deep learning in healthcare, however, has a set of issues, the main among them is patient data protection. Because healthcare companies are spread, and patient information is sensitive, traditional machine learning algorithms sometimes ask for the centralization of data, which poses challenges. Federated Learning has become recognized as a viable response to these issues. Federated Learning allows training machine learning models without exchanging and centralizing raw data.

In this case study, we have examined a unique strategy called Robust Privacy-Preserving Decentralized Federated Learning (RPDFL) proposed by Tian et al. [18] in the context of digital healthcare. Data silos," where healthcare institutions keep patient data in separate databases, frequently plague digital healthcare systems. Given that the data is scattered around several institutions, developing comprehensive and efficient machine-learning models is difficult due to these data silos. Data interoperability and cooperation are hampered by this fragmentation, making it challenging to fully utilize deep learning, which benefits from extensive and varied training datasets. Traditional centralized Federated Learning approaches also have their problems. These include communication bottlenecks between clients, the possibility of malicious servers attempting to deduce gradients, and single points of failure in the central server.

By grouping Federated Learning clients into a logical ring structure, RPDFL presents a novel strategy motivated by the widely utilized distributed computing Ring-AllReduce technique. This structure gives Each client a left and a right neighbor node. The fundamental idea is that each client should only communicate information with the proper neighbors to ensure synchronized global model updates and reduce vulnerability to malicious servers.

The strong privacy protection system of RPDFL is one of its key characteristics. It uses a threshold secret sharing mechanism, a cryptographic method to protect gradient privacy. This method maintains data security even if healthcare edge nodes gracefully leave during Federated Learning training. Clients exchange gradients from their local models. This sharing mechanism must preserve the privacy of each client's gradients. The remaining clients can complete the program uninterrupted if a client withdraws or disconnects, providing continual instruction. This privacy protection measure is essential to guarantee that sensitive patient data is kept private throughout the Federated Learning procedure.

Using their local datasets, RPDFL edge nodes in the healthcare industry produce local models across a number of training rounds. Stochastic Gradient Descent (SGD) is used to update these local models under the direction of the acquired global model. The Ring-Allreduce technique is used to broadcast gradients, which

stand in for the updates needed to the model parameters. Because of the threshold secret sharing technique, privacy is maintained during the whole procedure. Gradient averaging is used to produce new global models once gradients have been compiled and global model modifications have been made. For the upcoming training cycle, these upgraded global models will take the place of the local models. Up till the target level of model performance is attained, this iterative procedure is continued.Gradient averaging is used to generate new global models once gradients have been accumulated and global models have been modified. These improved global models will replace the local models in the future training cycle. This iterative method is repeated until the desired level of model performance is achieved.

For the performance evaluation of RPDFL on healthcare data, the UCI-HAR dataset from UCI was used, which captures daily human activities via smartphones and was evaluated using a fully connected network architecture. This capability enables RPDFL to effectively overcome information barriers between healthcare organizations, facilitating highly efficient decision-making in complex scenarios.

Thus, a new data-sharing scheme was implemented by updating the execution processes of the CRT in the threshold secret sharing scheme, allowing healthcare edges to drop out during training without causing data leakage and ensuring the robustness of the RPDFL training. The RPDFL scheme also supports edge dropout during the training process while preserving local gradient privacy. Security analysis demonstrates that RPDFL is highly secure under the HbC security model.

RPDFL exhibits promising performance in terms of privacy protection and accuracy. RPDFL obtains an accuracy level of around 98% by the fifth training round, which is an impressive accomplishment given its emphasis on privacy protection. Despite slower convergence after this first phase, RPDFL's accuracy is nevertheless on par with the conventional federated averaging (FedAvg) method. It demonstrates how RPDFL can maintain good performance while guarding against gradient privacy leaks. RPDFL has had a thorough security evaluation and has been shown to be extremely secure. According to security studies, RPDFL is resilient against harmful attacks and data breaches, protecting patient information's privacy and confidentiality.

RPDFL is a revolutionary technique for addressing the critical concerns of privacy and scalability in digital healthcare systems. RPDFL integrates a novel ring Federated Learning structure, a Ring-AllReduce-based data sharing mechanism, and a robust threshold secret sharing mechanism to provide a secure and effective solution for Federated Learning in the healthcare business. Figure 4.1 shows the data sharing process of Ring-AllReduce. Because of its extensive security features and promising performance, it is an excellent choice for healthcare organizations wishing to utilize the potential of deep learning while preserving patient privacy. As digital healthcare evolves, RPDFL is a valuable tool for harnessing the potential of machine learning in improving patient care and promoting medical research.

Then the performance of RPDFL in comparison to FedAvg and Gossip Learning is evaluated. Based on available information, Gossip Learning is the first implementation of decentralized learning. However, it is important to note that neither FedAvg nor Gossip Learning consider the privacy of gradients.

The training loss and testing accuracy for RPDFL, FedAvg, and Gossip Learning were recorded. Since FedAvg and Gossip Learning do not support edge dropout during the training process, RPDFL was also evaluated without edge dropout for comparison. It was observed that the accuracy of the RPDFL training model significantly surpasses that of Gossip Learning and is comparable to that of FedAvg. Additionally, RPDFL utilizes an enhanced CRT-based threshold secret sharing protocol to ensure gradient privacy. This protocol involves a truncation process, which results in some accuracy loss and a slower convergence speed. For instance, RPDFL converges slightly slower than FedAvg, but achieves an accuracy of 98% by the fifth round and maintains similar accuracy levels thereafter. Consequently, RPDFL offers excellent efficiency while safeguarding gradient privacy. This ensures that RPDFL can prevent malicious users from inferring the privacy of others in practical applications.

5. Open Research Challenges and Future Research Directions. In the rapidly evolving world of healthcare technology, data holds the key to ground-breaking advancements. This paper explores the dynamic world of Federated Learning in the healthcare domain, revealing the unique challenges posed by the distributed nature of medical data.

5.1. Non-IID Data Management. Data's inherent non-IID (Non-Independent and Identically Distributed) nature across various healthcare organizations presents substantial hurdles for Federated Learning in

the healthcare industry. Healthcare data is frequently gathered from several sources with variances in patient demographics, illness incidence, and treatment procedures, unlike typical centralized machine learning, where data is homogeneous and evenly disseminated. Using common Federated Learning techniques is challenging due to this heterogeneity directly. Innovative approaches like federated transfer learning and domain adaptation are being explored to tackle this issue [27]. These methods seek to account for the differences in data distributions while adapting and transferring information gained from one institution's data to another. Federated transfer learning makes it possible to share knowledge effectively by utilizing pre-trained models and fine-tuning them using data relevant to the institution.

5.2. Privacy-Preservation. The delicate nature of patient data needs sophisticated privacy-preserving strategies in Federated Learning since privacy is of the utmost significance in the healthcare industry. Traditional Federated Learning guarantees that data stays local to the institutions and is decentralized, but more developments are needed to improve privacy [28]. Differential privacy, a potential method, adds controlled noise to the model updates to secure patient information while allowing for useful learning. Homomorphic encryption is one of the most sophisticated encryption methods that can contribute to the protection of data throughout transmission and aggregation procedures. By implementing these privacy-enhancing strategies, Federated Learning may preserve patient confidentiality and adhere to stringent data protection laws.

5.3. Cross-Modal Data Fusion. The integration of multi-modal data develops as a crucial endeavor in the scene of healthcare improvement. It requires integrating many data types, from electronic health records to complex medical images, while respecting the fundamental principles of privacy protection. Modern approaches like federated transfer learning and meta-learning are simultaneously integrated, taking center stage [29]. By utilizing these methods to their full potential, the healthcare industry benefits from quick and seamless knowledge transfer, effective information translation and application across various contexts and activities, and strategic sharing of insights and expertise across institutions.

5.4. Interoperability and Data Harmonization. Integrating interoperability and data harmonization techniques is a crucial requirement in the healthcare field. These methods are essential for reducing inequities from disparate data formats and semantic details used by various healthcare facilities. Addressing the requirements for explainability and interpretability in the context of Federated Learning simultaneously becomes crucial. It requires developing models that produce reliable outcomes and foster mutual understanding and trust between patients and medical professionals. Model interpretability is a goal that has increased importance since it facilitates a broader understanding and acceptance of the results of Federated Learning projects.

5.5. Resource Management and Communication Efficiency. The successful deployment of Federated Learning in the healthcare sector depends on the effective allocation and use of resources. Notably, the need for effective communication is emphasized. It calls for developing novel strategies like decentralized aggregation and compressed model updates, which are key to reducing bandwidth requirements and boosting communication effectiveness. Healthcare systems can balance resource conservation and the seamless flow of vital information by carefully improving these strategies, eventually paving the way for a strong and effective Federated Learning framework.

The advancement of Federated Learning in healthcare, overcoming difficulties and supporting its effective application for improved patient care and medical insights [30]. By focusing on these future research routes, Federated Learning in healthcare may evolve further and contribute to better patient outcomes, customized treatment, and collaborative medical research as long as privacy, security, and ethical considerations are addressed.

6. Conclusion. Federated Learning in the healthcare industry has enormous potential to improve clinical decision-making, medical research, and customized treatment. Federated Learning makes it possible to use machine learning on decentralized healthcare data while protecting patient privacy by utilizing the strength of distributed computing and safe data sharing. This article has offered a comprehensive review of the current trends and issues facing the field of Federated Learning in healthcare. We explored how Federated Learning has emerged as a potential method for leveraging decentralized healthcare data while maintaining privacy and security. A case study on privacy preservation is presented in the context of digital healthcare. We have

also covered the key challenges that must be overcome, such as non-IID data management, interoperability, communication efficiency, and multi-model data integration. Federated Learning has a promising future in the medical field. Medical professionals and researchers should keep working collectively to create novel approaches, improve current techniques, and solve the legal and ethical challenges involving the sharing of medical data. To guarantee that Federated Learning helps the healthcare business and the persons it serves, it is critical to prioritize its appropriate and ethical deployment.

REFERENCES

[1] Choudhury, O., Gkoulalas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G. & Das, A. Differential Privacy-enabled Federated Learning for Sensitive Health Data. (2020)

[2] Shailaja, K., Seetharamulu, B. & Jabbar, M. Machine Learning in Healthcare: A Review. *2018 Second International Conference On Electronics, Communication And Aerospace Technology (ICECA)*. pp. 910-914 (2018)

[3] Yang, Q., Liu, Y., Chen, T. & Tong, Y. Federated Machine Learning: Concept and Applications. (Association for Computing Machinery,2019), https://doi.org/10.1145/3298981

[4] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H., Albarqouni, S., Bakas, S., Galtier, M., Landman, B., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R., Trask, A., Xu, D., Baust, M. & Cardoso, M. The future of digital health with federated learning. *Npj Digital Medicine*. **3** (2020,12)

[5] Li, J., Meng, Y., Ma, L., Du, S., Zhu, H., Pei, Q. & Shen, X. A Federated Learning Based Privacy-Preserving Smart Healthcare System. *IEEE Transactions On Industrial Informatics*. **18**, 2021-2031 (2022)

[6] Abdulrahman, S., Tout, H., Ould-Slimane, H., Mourad, A., Talhi, C. & Guizani, M. A Survey on Federated Learning: The Journey From Centralized to Distributed On-Site Learning and Beyond. *IEEE Internet Of Things Journal*. **PP** (2020,10)

[7] Prayitno, Shyu, C., Putra, K., Chen, H., Tsai, Y., Hossain, K., Jiang, W. & Shae, Z. A Systematic Review of Federated Learning in the Healthcare Area: From the Perspective of Data Properties and Applications. *Applied Sciences*. **11** (2021), https://www.mdpi.com/2076-3417/11/23/11191

[8] Gupta, D., Kayode, O., Bhatt, S., Gupta, M. & Tosun, A. Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare. (2021,11)

[9] Xu, J., Glicksberg, B., Su, C., Walker, P., Bian, J. & Wang, F. Federated Learning for Healthcare Informatics. *Journal Of Healthcare Informatics Research*. **5** pp. 1-19 (2021,3)

[10] Stripelis, D., Ambite, J., Lam, P. & Thompson, P. Scaling Neuroscience Research using Federated Learning. (2021,2)

[11] Choudhury, O., Gkoulalas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G. & Das, A. Differential Privacy-enabled Federated Learning for Sensitive Health Data. *ArXiv*. **abs/1910.02578** (2019)

[12] Liu, Z., Chen, Y., Zhao, Y., Yu, H., Liu, Y., Bao, R., Jiang, J., Nie, Z., Xu, Q. & Yang, Q. Contribution-aware federated learning for smart healthcare. *Proceedings Of The AAAI Conference On Artificial Intelligence*. **36**, 12396-12404 (2022)

[13] Sun, L. & Wu, J. A Scalable and Transferable Federated Learning System for Classifying Healthcare Sensor Data. *IEEE Journal Of Biomedical And Health Informatics*. **27**, 866-877 (2023)

[14] Patel, V., Bhattacharya, P., Tanwar, S., Gupta, R., Sharma, G., Bokoro, P. & Sharma, R. Adoption of federated learning for healthcare informatics: Emerging applications and future direction. *IEEE Access*. (2022)

[15] Wu, Q., Chen, X., Zhou, Z. & Zhang, J. FedHome: Cloud-Edge Based Personalized Federated Learning for In-Home Health Monitoring. *IEEE Transactions On Mobile Computing*. **21**, 2818-2832 (2022,8)

[16] Lu, W., Wang, J., Chen, Y., Qin, X., Xu, R., Dimitriadis, D. & Qin, T. Personalized federated learning with adaptive batchnorm for healthcare. *IEEE Transactions On Big Data*. (2022)

[17] Islam, T., Ghasemi, R. & Mohammed, N. Privacy-preserving federated learning model for healthcare data. *2022 IEEE 12th Annual Computing And Communication Workshop And Conference (CCWC)*. pp. 0281-0287 (2022)

[18] Tian, Y., Wang, S., Xiong, J., Bi, R., Zhou, Z. & Bhuiyan, M. Robust and Privacy-Preserving Decentralized Deep Federated Learning Training: Focusing on Digital Healthcare Applications. *IEEE/ACM Transactions On Computational Biology And Bioinformatics*. pp. 1-12 (2023)

[19] Kaur, J., Verma, R., Alharbe, N., Agrawal, A. & Khan, P. Importance of Fog Computing in Healthcare 4.0. (2020,8)

[20] Razdan, S., & Sharma, S. (2022). Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies. *IETE Technical Review*, 39(4), 775–788. Taylor & Francis. doi:10.1080/02564602.2021.1927863.

[21] Mohanta, B., Das, P. & Patnaik, S. Healthcare 5.0: A Paradigm Shift in Digital Healthcare System Using Artificial Intelligence, IOT and 5G Communication. *2019 International Conference On Applied Machine Learning (ICAML)*. pp. 191-196 (2019)

[22] Vishnu, S., Ramson, S. & Jegan, R. Internet of Medical Things (IoMT) - An overview. *2020 5th International Conference On Devices, Circuits And Systems (ICDCS)*. pp. 101-104 (2020)

[23] Pournik, O., Ghalichi, L., Gallos, P. & Arvanitis, T. The Internet of Medical Things: Opportunities, Benefits, Challenges and Concerns. *Studies In Health Technology And Informatics*. **309** (2023,10)

[24] Dhade, P. & Shirke, P. Federated Learning for Healthcare: A Comprehensive Review. *Engineering Proceedings*. **59** (2023), https://www.mdpi.com/2673-4591/59/1/230

[25] Prayitno, Shyu, C., Putra, K., Chen, H., Tsai, Y., Hossain, K., Jiang, W. & Shae, Z. A Systematic Review of Federated Learning in the Healthcare Area: From the Perspective of Data Properties and Applications. *Applied Sciences*. **11** (2021), https://www.mdpi.com/2076-3417/11/23/11191

[26] De Falco, I., Della Cioppa, A., Koutny, T., Ubl, M., Krcma, M., Scafuri, U. & Tarantino, E. A Federated Learning-

Inspired Evolutionary Algorithm: Application to Glucose Prediction. *Sensors*. **23** (2023), https://www.mdpi.com/1424-8220/23/6/2957

[27] Lu, Z., Pan, H., Dai, Y., Si, X. & Zhang, Y. Federated Learning With Non-IID Data: A Survey. *IEEE Internet Of Things Journal*. **11**, 19188-19209 (2024)

[28] Maurya, J. & Prakash, S. Privacy Preservation in Federated Learning: its Attacks and Defenses. *2023 3rd International Conference On Pervasive Computing And Social Networking (ICPCSN)*. pp. 1042-1047 (2023)

[29] Ahmed, S., Alam, M., Afrin, S., Rafa, S., Rafa, N. & Gandomi, A. Insights into Internet of Medical Things (IoMT): Data fusion, security issues and potential solutions. *Information Fusion*. **102** pp. 102060 (2024), https://www.sciencedirect.com/science/article/pii/S1566253523003767

[30] Muazu, T., Yingchi, M., Muhammad, A., Ibrahim, M., Samuel, O. & Tiwari, P. IoMT: A Medical Resource Management System Using Edge Empowered Blockchain Federated Learning. *IEEE Transactions On Network And Service Management*. **21**, 517-534 (2024)

[31] Dai, H., Imran, M. & Haider, N. Blockchain-Enabled Internet of Medical Things to Combat COVID-19. *IEEE Internet Of Things Magazine*. **3**, 52-57 (2020)

[32] Rani, S., Kataria, A., Kumar, S. & Tiwari, P. Federated learning for secure IoMT-applications in smart healthcare systems: A comprehensive review. *Knowledge-Based Systems*. **274** pp. 110658 (2023), https://www.sciencedirect.com/science/article/pii/S0950705123004082

[33] Alamleh, A., Albahri, O., Zaidan, A., Albahri, A., Alamoodi, A., Zaidan, B., Qahtan, S., Alsatar, H., Al-Samarraay, M. & Jasim, A. Federated Learning for IoMT Applications: A Standardization and Benchmarking Framework of Intrusion Detection Systems. *IEEE Journal Of Biomedical And Health Informatics*. **27**, 878-887 (2023)

[34] Ali, M., Tariq, M., Naeem, F. & Kaddoum, G. Federated Learning for Privacy Preservation in Smart Healthcare Systems: A Comprehensive Survey. *IEEE Journal Of Biomedical And Health Informatics*. **PP** (2022,6)

[35] Arya, M. & G, H. Ensemble Federated Learning for Classifying IoMT Data Streams. *2022 IEEE 7th International Conference For Convergence In Technology (I2CT)*. pp. 1-5 (2022)

[36] Lin, H., Han, J., Wu, P., Zhu, L., Wang, J. & Tu, J. Machine Learning and Human-Machine Trust in Healthcare: A Systematic Survey. *SSRN Electronic Journal*. (2023), https://api.semanticscholar.org/CorpusID:257704958

[37] Durga, S., Nag, R. & Daniel, E. Survey on Machine Learning and Deep Learning Algorithms used in Internet of Things (IoT) Healthcare. *2019 3rd International Conference On Computing Methodologies And Communication (ICCMC)*. pp. 1018-1022 (2019)

[38] Al-Dhief, F., Latiff, N., Malik, N., Salim, N., Baki, M., Albadr, M. & Mohammed, M. A Survey of Voice Pathology Surveillance Systems Based on Internet of Things and Machine Learning Algorithms. *IEEE Access*. **8** pp. 64514-64533 (2020)

Edited by: Katarzyna Wasielewska-Michniewska
Review paper
Received: Nov 23, 2023
Accepted: Jul 29, 2024