Journal of Theoretical and Applied Information Technology

15th November 2023. Vol.101. No 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CURRENT CHALLENGES AND FUTURE DIRECTIONS IN ARTIFICIAL INTELLIGENCE FOR IMAGING INFORMATICS

J. RAVINDRA BABU[1], BHARGAVI PEDDI REDDY[2], VANGIPURAM SESHA SRINIVAS[3], A. L. SREENIVASULU[4], K V S S RAMAKRISHNA[5], *D N V SATYANARAYANA[6], C.D.VARAPRASAD[7]*

[1]Associate Professor, Department of ECE, P. V. P. Siddhartha institute of Technology, Vijayawada, Andhra Pradesh, India. E-mail: jrb0009@gmail.com

[2]Associate Professor, Department of CSE, Vasavi College of Engineering, Hyderabad, India
E-mail: bhargavi@staff.vce.ac.in

[3]Associate Professor, Department of IT, R.V.R. & J. C. College of Engineering, Guntur, Andhra Pradesh, India. E-mail: vangipuramseshu@gmail.com

[4]Professor, Department of CSE, Vignana Bharathi Institute of Technology, Ghatkesar, Medchal, Hyderabad, Telangana, India. E-mail: akula.srinivasulu@vbithyd.ac.in

[5]Department of IT, Vignan's Nirula Institute of Technology and Science for Women, Guntur, Andhra Pradesh, India. Email: ksai.mb@gmail.com

[6]Associate Professor, Department of Chemical Engineering, R.V.R & J. C. College of Engineering, Guntur, Andhra Pradesh, India. E-mail: dnvsatya2001@gmail.com

[7]Senior lecturer, School of Electrical and Computer Engineering, Haramaya Institute of Technology, Haramaya University, DireDawa, Ethiopia. Email id: varaprasadchitte@gmail.com

ABSTRACT

Artificial Intelligence (AI) has made significant strides in healthcare, revolutionizing various aspects of medical diagnosis, treatment, and patient care. However, the adoption of AI in medicine is hindered by challenges related to model interpretability, generalization across different healthcare domains, and data privacy concerns. This research paper explores the concepts of explainable AI (XAI), domain adaptation, and federated learning in the context of healthcare, and their potential to address these challenges. Discusses the significance of developing AI models that are explainable, adaptable across different domains, and capable of leveraging distributed data sources through federated learning to enhance medical decision-making while maintaining patient privacy.

Keywords: *Artificial Intelligence, Explainable AI, Federated Learning, Medical.*

1. INTRODUCTION

Artificial Intelligence (AI) has emerged as a powerful tool in the field of medicine, with the potential to transform healthcare delivery, diagnosis, and treatment. AI techniques, such as machine learning and deep learning, enable computers to learn from large amounts of data and make predictions or decisions with remarkable accuracy. However, the successful integration of AI into healthcare faces several challenges that must be addressed to ensure its widespread adoption and acceptance [1]. One of the primary challenges is the lack of interpretability in AI models. In medical applications, it is crucial to understand the rationale behind AI predictions or recommendations to build trust and facilitate clinical decision-making. Explainable AI (XAI) techniques aim to provide transparent and interpretable models, enabling healthcare professionals to understand how AI arrives at its conclusions. XAI techniques not only enhance the trustworthiness of AI systems but also facilitate regulatory compliance and ethical considerations in healthcare. Another challenge is the limited generalization of AI models across different healthcare domains [2]. Medical data is diverse and heterogeneous, varying across specialties, institutions, and patient populations [3]. AI models trained on data from one domain often struggle to perform well when applied to a different domain due to distributional differences. Domain adaptation techniques seek to bridge this gap by adapting models to new target domains, leveraging knowledge from source domains while accounting for domain shifts. Domain-adaptive AI can enhance the transferability and robustness of AI models,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

making them more applicable in real-world medical scenarios [4].

Furthermore, data privacy and security concerns pose significant barriers to AI adoption in healthcare. Patient data, including electronic health records (EHRs) and medical images, are sensitive and subject to privacy regulations. Sharing such data centrally for model training raises privacy risks [5]. Federated Learning (FL) offers a solution by enabling distributed model training across multiple data sources while keeping data locally stored and secure. FL allows collaborative learning without data sharing, preserving patient privacy and addressing legal and ethical concerns [6].

This research paper aims to explore the concepts of explainable AI, domain adaptation, and federated learning in the context of medicine. It examines the significance of developing AI models that are explainable, adaptable across different healthcare domains, and capable of leveraging distributed data sources through federated learning. By addressing these challenges, AI in medicine can become more transparent, versatile, and privacy-preserving, ultimately improving patient outcomes and enabling evidence-based decision-making [7]. The subsequent sections of this paper will delve into the intricacies of explainable AI, domain adaptation, and federated learning in healthcare, highlighting their techniques, applications, synergies, and future directions. Additionally, case studies and evaluations will be presented to demonstrate the effectiveness of these approaches in enhancing medical decision-making and patient care [8]. Finally, the paper will discuss the remaining challenges and provide recommendations for future research and implementation to further advance the field of explainable, domain-adaptive, and federated AI in medicine.

2. EXPLAINABLE AI (XAI) IN MEDICINE

In the field of medicine, the interpretability of AI models is of paramount importance. Healthcare professionals and patients need to understand the reasoning behind AI-generated predictions or decisions to trust and effectively utilize these systems. Interpretability fosters transparency, accountability, and the ability to validate and refine models in the medical domain [9]. Moreover, regulatory bodies often require explanations for AI-based decisions to ensure compliance with ethical and legal standards. Several techniques have been developed to enhance the explainability of AI models in medical applications [10]:

Rule-based methods generate interpretable decision rules that mimic human reasoning. These methods use expert knowledge or learn rules from the data itself. Rule-based models, such as decision trees and rule lists, offer a clear understanding of how input features contribute to the final decision.

Local explanation methods focus on explaining individual predictions or decisions. They highlight the specific features or factors that influenced the model's output for a particular case. Techniques like LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) provide local interpretability by assigning feature importance scores.

Global explanation methods aim to provide an understanding of the overall behavior and decision-making process of AI models. They reveal the most influential features and their relationships, enabling a holistic view of the model's decision strategy. Methods such as feature importance analysis, surrogate models, and Partial Dependence Plots (PDP) fall into this category.

Model visualization techniques employ visual representations to convey AI model behavior. Visualizations can provide insights into model internals, such as feature interactions, decision boundaries, or learned representations. Techniques like heatmaps, saliency maps, and activation maximization can help interpret deep learning models in medical imaging and pathology.

Certainty and confidence measures quantify the model's confidence in its predictions. They can be used to identify cases where the model's output may be less reliable or uncertain. These measures allow healthcare professionals to assess the trustworthiness of AI predictions and decide when to seek additional confirmatory tests or expert opinions.

Table 1 describes XAI in deep learning models applied to clinical problems. We highlight that CNNs are the most extensively used deep learning models, and gradient-based explainability is the most widely employed principle. The amount of explainability is usually local [11] – [16].

TABLE 1: XAI Models Used in Clinical Applications: A Summary

Application	XAI / Explainability Level	Prediction Model
Preclinical Relevance Assessment	Model combination/G	GNN, FNN
Detection And Prediction For Alzheimer's Disease	SHAP / L & G	RF
Deterioration Risk Prediction Of Hepatitis	SHAP, LIME, PDP / L & G	RF
Non-Communicable Diseases Prediction Deep	SHAP / L & G	DNN
Spinal Posture Classification	LIME / L	SVM
Classification of Estrogen Receptor Status	Smooth Grad / L	DCNN
Differential Diagnosis Of COVID-19	Grad-CAM / G	CNN
Prediction of Mortality	Shapley additive / L	RNN
Medical Image Segmentation	Combination / L & G CNN	Attention Mechanism
COVID-19 Diagnosis	Grad-CAM ++, LRP / L	DNN

L: Local; G: Global; ANN: Artificial Neural Network; GNN: Graph Neural Networks; FNN: Fuzzy Neural Network; DNN: Deep Neural Networks; RNN: Recurrent Neural Network; FCN: Fully Convolutional Network; DCNN: Deep Convolutional Neural Network; VGG-16: GG-Very-Deep-16 CNN; VGG-19: GG-Very-Deep-19 CNN; Fuzzy-enhanced CNN: Fuzzy-enhanced Convolutional Neural Network; RF: Random forest; SVM: Support Vector Machine; MLP: Muti-Layer Perceptron; LSTM: Long Short-Term Memory network; PDP: Partial Dependence Plot; t-SNE: t-distributed Stochastic Neighbor Embedding. Seven Network: SqueezeNet, Inception, ResNet, ResNeXt, Xception, ShuffleNet, DenseNet; Five XAI methods: Saliency, guided backpropagation, integrated gradients, input gradients, and DeepLIFT.

With high-performance deep learning models of increasing complexity, the way specific results were obtained is not accessible to humans. A number of methods were proposed to increase explainability in recent years and a selection of which is listed in Table 1 and these techniques are shown in figure 1.

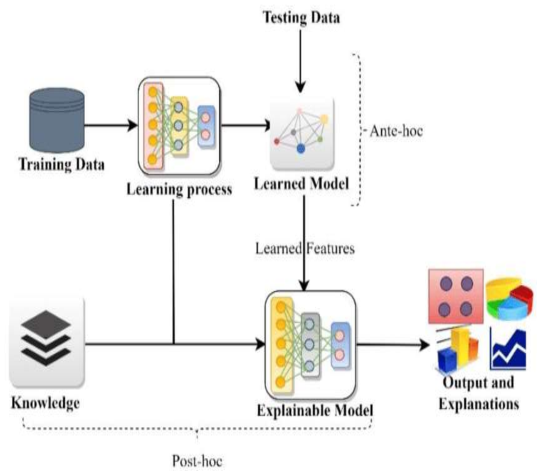

Figure 1: A flowchart example for an explainable artificial intelligence model

2.1 Balancing Transparency and Performance in Healthcare AI Models

While interpretability is crucial, it is also essential to strike a balance between model transparency and performance. Highly interpretable models, such as simple rule-based systems, may sacrifice accuracy for transparency. On the other hand, more complex models like deep neural networks may offer superior performance but are inherently less interpretable. Research in XAI focuses on developing methods that optimize this trade-off, allowing clinicians to trust the models' predictions while maintaining competitive performance levels.

2.2 Challenges in XAI Adoption in Medicine

The adoption of XAI in medicine faces several challenges:

Medical data is often complex, high-dimensional, and multi-modal, including clinical notes, imaging data, genetic information, and time-series data. Extracting meaningful and easily understandable explanations from such data poses unique challenges.

Deep learning models, with their complex architectures and millions of parameters, are challenging to interpret. Techniques such as layer-wise relevance propagation (LRP) and Grad-CAM (Gradient-weighted Class Activation

Mapping) aim to address this issue by attributing importance to different parts of the input data.

Explainability requirements imposed by regulatory bodies, such as the General Data Protection Regulation (GDPR) and the Health Insurance

Portability and Accountability Act (HIPAA), pose challenges for AI implementation. Developing XAI techniques that satisfy regulatory demands without compromising performance is a critical task.

As models become more complex and incorporate larger amounts of data, achieving both high interpretability and accuracy becomes more challenging. Striking the right balance is crucial to ensure effective clinical adoption and utilization of AI systems.

Explainable AI techniques play a vital role in ensuring the trustworthiness and transparency of AI models in medicine. By providing interpretable explanations, clinicians can understand and validate AI-generated decisions, leading to improved patient care and informed decision-making. Overcoming the challenges associated with XAI adoption in medicine will pave the way for the responsible and effective integration of AI into clinical practice.

3. DOMAIN ADAPTATION IN HEALTHCARE

3.1 The Need for Adaptable AI Models across Different Healthcare Domains

Healthcare data is diverse and collected from various sources such as hospitals, clinics, research institutions, and wearable devices. These data sources often exhibit variations in terms of patient demographics, clinical protocols, data collection methodologies, and disease prevalence. Consequently, AI models trained on data from one healthcare domain may struggle to generalize and perform well when applied to different domains. Therefore, there is a need for domain adaptation techniques to make AI models adaptable across different healthcare domains.

3.2 Challenges in Domain Adaptation for Medical Applications

Domain adaptation in healthcare encounters specific challenges:

Data collected from different healthcare domains may have distinct distributions due to variations in patient populations, clinical practices, or data collection protocols. These distribution shifts can lead to a significant performance drop when models trained on one domain are applied to another. Overcoming the distribution shift is a key challenge in domain adaptation.

Labeled data in healthcare is often scarce and expensive to obtain due to the need for expert annotations and privacy concerns. In the context of domain adaptation, acquiring labeled data from the target domain can be particularly challenging, as it may require additional resources and expertise.

Healthcare data comprises various modalities, such as electronic health records (EHRs), medical images, genetic data, and sensor data. Integrating and adapting models across these different modalities introduces additional complexity in domain adaptation.

3.3 Techniques for Domain Adaptation in Medical AI

Several techniques have been developed to address domain adaptation challenges in medical AI:

Feature-based methods aim to align the feature representations between the source and target domains. This can be achieved through techniques such as domain adversarial neural networks (DANN) and maximum mean discrepancy (MMD). These methods encourage the model to learn domain-invariant features that generalize well across domains.

Instance-based methods focus on adapting the model's predictions on a per-instance basis. This can be done through techniques like self-training, where the model iteratively generates pseudo-labels for unlabeled target domain data and includes them in the training process. This approach leverages the unlabeled target domain data to refine the model's predictions.

Model-based methods involve training a single model that can effectively generalize to multiple domains. This can be achieved through approaches like multi-task learning, where the model is trained to perform well on multiple related tasks simultaneously, leveraging the shared knowledge across domains.

Transfer learning techniques, such as fine-tuning pre-trained models, can be utilized in domain adaptation. Pre-training models on a large-scale dataset can capture general knowledge that can be adapted to specific healthcare domains with limited labeled data.

3.4 Evaluation and Validation of Domain Adaptation Techniques

The evaluation of domain adaptation techniques in healthcare requires careful consideration. It is important to assess the performance of adapted models on the target domain while comparing them against models trained only on the source domain or models trained from scratch on the target domain. Evaluation metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) are commonly used to assess model performance across different domains.

3.5 Challenges and Future Directions

Domain adaptation in healthcare is a complex and ongoing research area. Several challenges and future directions include:

Domain adaptation techniques should be robust to handling data that deviates significantly from the known domains. AI models should be able to identify and handle out-of-distribution samples effectively.

Developing unsupervised domain adaptation methods that eliminate the need for labeled target domain data would be highly valuable. This would enable adaptation to new domains without requiring expensive annotations.

As healthcare data often consists of multiple modalities, developing domain adaptation techniques that can adapt models across heterogeneous data sources is crucial for comprehensive AI applications in healthcare.

Domain-adaptive AI models need to be rigorously validated in clinical settings to ensure their safety and efficacy. Additionally, ethical considerations, such as bias mitigation and fairness, should be addressed in the adaptation process to avoid exacerbating existing disparities.

Domain adaptation techniques are essential for making AI models adaptable across different healthcare domains. Overcoming challenges related to data distribution shifts, limited labeled data, and heterogeneous data modalities can enable the development of robust and generalizable AI systems in medicine. Continued research and innovation in this field will facilitate the effective utilization of AI in various healthcare settings while improving patient care and outcomes.

4. FEDERATED LEARNING FOR PRIVACY-PRESERVING MEDICAL AI

4.1 Privacy Challenges in Medical AI

Medical data contains sensitive and personal information, making privacy a significant concern in the development and deployment of AI models. Traditional approaches involve centralizing the data for training, which raises privacy and security risks. Patients' privacy rights, legal regulations (e.g., HIPAA), and ethical considerations necessitate the exploration of privacy-preserving techniques. Federated Learning (FL) offers a promising solution by enabling collaborative model training while keeping the data decentralized and secure.

4.2 Understanding Federated Learning

Federated Learning is a distributed machine learning paradigm that allows model training across multiple edge devices or data sources without centrally aggregating the data. Instead, the model is sent to the data sources, and local updates are performed on the data, preserving privacy. Only the model updates, instead of raw data, are exchanged between the data sources and a central server. The central server aggregates these updates to create an improved global model that benefits from the collective knowledge without directly accessing the private data.

4.3 Advantages of Federated Learning in Healthcare

Federated Learning offers several advantages for privacy-preserving medical AI:

FL eliminates the need to transmit sensitive patient data to a central location, reducing the risk of data breaches and unauthorized access. Patient data remains under local control, ensuring compliance with privacy regulations and enhancing patient trust.

In healthcare, data is often collected from diverse sources, such as hospitals, clinics, and wearable devices. FL accommodates this heterogeneity by allowing training on data sources with different distributions and characteristics without requiring data sharing.

FL enables large-scale collaboration and data sharing among healthcare institutions, fostering research and innovation. It facilitates the inclusion of smaller clinics or institutions that may have limited resources or data, enabling them to participate in model training and benefit from the collective knowledge.

FL supports training models directly on edge devices, such as smartphones or wearable devices, which are becoming increasingly powerful. This enables personalized and real-time AI applications at the edge, enhancing patient care and health monitoring.

4.4 Challenges and Considerations in Federated Learning for Healthcare

While FL offers promising benefits, it also presents challenges and considerations:

FL requires frequent communication between the central server and data sources to exchange model updates. Optimizing communication protocols and reducing the bandwidth requirements are essential for efficient FL implementation, especially in resource-constrained healthcare environments.

Healthcare data from different sources may exhibit variations in terms of data quality, noise, or biases. Ensuring model convergence and generalization across heterogeneous data sources remains a challenge in FL.

FL should incorporate privacy-preserving techniques, such as secure aggregation, differential privacy, and encryption, to further safeguard the privacy of sensitive medical data. These techniques ensure that even in the presence of malicious participants, individual data remains protected.

FL must address ethical considerations, such as fairness, bias, and transparency. Bias in local data sources can impact the global model, and mechanisms should be in place to mitigate and monitor such biases. Additionally, ensuring transparency and interpretability of FL models is crucial to maintain trust and accountability.

4.5 Real-world Applications of Federated Learning in Healthcare

Federated Learning has promising applications in healthcare:

FL can enable the development of robust clinical decision support systems by training models across multiple healthcare institutions, capturing diverse patient populations and clinical practices.

FL allows for the collective training of AI models on data from different geographic regions or patient cohorts, enabling more accurate disease prediction and risk stratification models.

FL facilitates secure and privacy-preserving data sharing for collaborative research, enabling large-scale analysis while preserving patient privacy.

FL can support personalized medicine by training models on local data sources, such as wearable devices, enabling personalized health monitoring and interventions while preserving data privacy.

Federated Learning provides a privacy-preserving approach for training AI models in healthcare. By allowing collaborative model training without sharing sensitive patient data, FL enables the development of robust and scalable medical AI solutions. Addressing the challenges and considerations associated with FL in healthcare will pave the way for its wider adoption, fostering innovation, and improving patient care while respecting privacy and security requirements.

5. INTEGRATION OF XAI, DOMAIN ADAPTATION, AND FEDERATED LEARNING

The integration of Explainable AI (XAI), Domain Adaptation, and Federated Learning (FL) in medical applications can lead to powerful and privacy-preserving AI systems with enhanced interpretability and adaptability. Each of these techniques addresses specific challenges in medical AI, and their combination can provide comprehensive solutions that benefit healthcare professionals and patients.

The integration of XAI with FL and domain adaptation techniques can enhance the interpretability of models trained on decentralized and diverse data sources. By providing explanations for model predictions or decisions in a federated setting, healthcare professionals can gain insights into the reasoning behind AI-generated recommendations and trust the system's output. XAI techniques, such as local explanations and model visualization, can help identify features or patterns contributing to predictions across different domains, thereby improving transparency and facilitating model validation.

Domain adaptation techniques integrated with FL can address the challenge of data distribution shift when training models on decentralized and diverse data sources. By adapting models to different healthcare domains, FL can create models that generalize well across various institutions, patient populations, and data modalities. This integration allows AI models to capture domain-specific knowledge while leveraging the collective intelligence of multiple sources. The adaptation process can help align the models' representations and decision boundaries to improve performance in target domains, enabling more effective and generalizable AI systems in healthcare.

Integrating privacy-preserving XAI techniques with FL and domain adaptation ensures that sensitive patient data remains protected while providing interpretable AI models. Privacy-enhancing techniques, such as federated explainability methods and secure local model interpretation, enable clinicians to understand the model's decisions without compromising patient privacy. By providing local and global explanations in a privacy-preserving manner, healthcare professionals can trust the AI models' recommendations without direct access to patient-specific data.

5.1 Real-World Applications and Benefits

The integration of XAI, domain adaptation, and FL has several real-world applications and benefits in healthcare:

By combining domain adaptation and FL, AI models can be trained on diverse patient populations, leading to personalized treatment recommendations that account for variations in demographics, clinical protocols, and disease prevalence across different domains. XAI techniques can provide transparent explanations for these personalized recommendations, fostering trust and facilitating shared decision-making between clinicians and patients.

FL allows the collaboration of multiple healthcare institutions for disease surveillance and outbreak detection without sharing sensitive patient data. Domain adaptation techniques can enable models to adapt to different geographical regions or evolving disease patterns. XAI provides interpretability to these surveillance models, enabling public health officials to understand and act upon the predictions effectively.

The integration of XAI, domain adaptation, and FL can support remote monitoring and telemedicine applications. FL enables the training of models on data from wearable devices or remote sensors while preserving data privacy. Domain adaptation techniques ensure the models can adapt to individual patients' characteristics. XAI techniques provide explanations for the models' predictions, empowering patients and healthcare professionals to understand and trust the remote monitoring systems.

Further research is needed to develop privacy-preserving XAI methods that can explain AI models' decisions without compromising patient privacy in federated and domain adaptive settings. Techniques such as secure aggregation of local explanations and privacy-aware visualization can be explored.

Future work should focus on addressing bias and ensuring fairness in the integrated framework. By considering ethical considerations and incorporating fairness metrics during domain adaptation and FL, the resulting AI systems can mitigate biases and promote equitable healthcare outcomes.

Establishing standards and evaluation frameworks for integrated XAI, domain adaptation, and FL approaches in healthcare is essential. Standardization can promote interoperability, facilitate the sharing of best practices, and ensure the reliability and reproducibility of AI systems across different healthcare domains.

Successful integration of XAI, domain adaptation, and FL requires user-centric design and collaboration with healthcare professionals. Understanding their needs, preferences, and workflows is crucial for designing effective and usable AI systems in real-world clinical settings.

The integration of XAI, domain adaptation, and FL in medical AI offers synergistic benefits, including enhanced interpretability, generalizability, and privacy preservation. This integrated approach can lead to improved patient care, personalized treatment recommendations, collaborative disease surveillance, and remote monitoring applications. Addressing privacy concerns, bias, fairness, and standardization challenges will pave the way for the responsible and effective integration of these techniques in healthcare.

FL updates the models to achieve high performance by considering co-modeling step between different data sources (or client). FL is divided into the following categories shown in figure 2 and figure 3. Horizontal federated learning (HFL) is a system in which all the parties share the same feature space. Vertical federated learning (VFL) lets multiple parties that possess different attributes (e.g., features and/or labels) of the same data entity (e.g., a person) to jointly train a model.

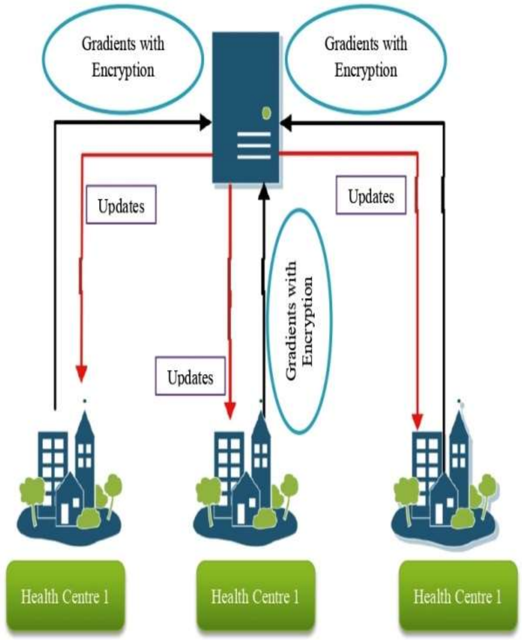

Figure 2: Horizontal Federate Learning

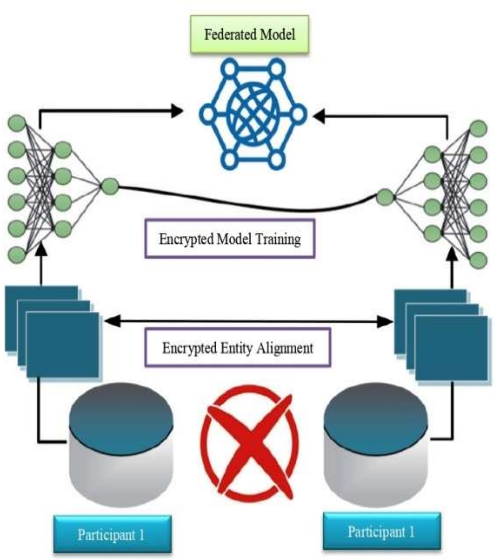

Figure 3: Vertical Federated Learning

6. CONCLUSION

The integration of explainable artificial intelligence (XAI), domain adaptation, and federated learning (FL) in the field of medicine presents a significant opportunity for the advancement of artificial intelligence used in healthcare systems. Each method tackles a different set of problems, and when they are integrated, they produce comprehensive solutions that improve interpretability, flexibility, and privacy protection. XAI methods make it possible for medical practitioners to comprehend the logic that behind AI-generated recommendations and so increase their level of confidence in the system. Methods of domain adaptation ensure that artificial intelligence models generalize well across various healthcare domains by taking into account variances in patient populations, data sources, and clinical practices. FL enables the development of medical AI solutions that are both resilient and scalable. This is made possible by the fact that FL enables collaborative model training while maintaining data privacy. The combination of XAI, domain adaptability, and FL offers a number of practical applications, including the enhancement of disease identification and risk classification, the facilitation of tailored medication, and the support of joint research and remote monitoring. These applications have the potential to improve patient care, make it possible to make decisions based on data, and make disease tracking easier.

REFERENCES

[1]. Giovanni Briganti and Olivier Le Moine. 2020. "Artificial Intelligence in Medicine: Today and Tomorrow," Frontiers in Medicine, 7.

[2]. Sajid Ali, Tamer Abuhmed, Shaker El-Sappagh, Khan Muhammad, Jose M. Alonso-Moral, Roberto Confalonieri, Riccardo Guidotti, Javier Del Ser, Natalia Díaz-Rodríguez, Francisco Herrera, "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence," Information Fusion, 2023.

[3]. Dash, S., Shakyawar, S.K., Sharma, M. et al. "Big data in healthcare: management, analysis and future prospects," J. Big Data 6, 54 (2019).

[4]. Kelly, C.J., Karthikesalingam, A., Suleyman, M. et al. "Key challenges for delivering clinical impact with artificial intelligence," BMC Med 17, 195 (2019).

[5]. Murdoch, B. "Privacy and artificial intelligence: challenges for protecting health information in a new era," BMC Med Ethics 22, 122 (2021).

[6]. Li Li, Yuxi Fan, Mike Tse, Kuo-Yi Lin, “A review of applications in federated learning,” Computers & Industrial Engineering, Volume 149, 2020.

[7]. A. Chaddad et al., "Explainable, Domain-Adaptive, and Federated Artificial Intelligence in Medicine," in IEEE/CAA Journal of Automatica Sinica, vol. 10, no. 4, pp. 859-876, April 2023.

[8]. Loh HW, Ooi CP, Seoni S, Barua PD, Molinari F, Acharya UR. “Application of xplainable artificial intelligence for healthcare: A systematic review of the last decade (2011-2022),” Comput Methods Programs Biomed. 2022 Nov; 226:107161, Epub 2022 Sep 27. PMID: 36228495.

[9]. Bas H.M. van der Velden, Hugo J. Kuijf, Kenneth G.A. Gilhuijs, Max A. Viergever, “Explainable artificial intelligence (XAI) in deep learning-based medical image analysis,” Medical Image Analysis, Volume 79, 2022.

[10]. Yu-Cheng Wang, Tin-Chih Toly Chen, Min-Chi Chiu, “An improved explainable artificial intelligence tool in healthcare for hospital recommendation,” Healthcare Analytics, Volume 3, 2023.

[11]. J. Jiménez-Luna, M. Skalic, N. Weskamp, and G. Schneider, “Coloring molecules with explainable artificial intelligence for preclinical relevance assessment,” J. Chem. Inf. Model., vol. 61, no. 3, pp. 1083–1094, Feb. 2021.

[12]. S. El-Sappagh, J. M. Alonso, S. M. R. Islam, A. M. Sultan, and K. S. Kwak, “A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer’s disease,” Sci. Rep., vol. 11, no. 1, p. 2660, Jan. 2021.

[13]. J. F. Peng, K. Q. Zou, M. Zhou, Y. Teng, X. Y. Zhu, F. F. Zhang, and J. Xu, “An explainable artificial intelligence framework for the deterioration risk prediction of hepatitis patients,” J. Med. Syst., vol. 45, no. 5, p. 61, Apr. 2021.

[14]. K. Davagdorj, J. W. Bae, V. H. Pham, N. Theera-Umpon, and K. H. Ryu, “Explainable artificial intelligence based framework for noncommunicable diseases prediction,” IEEE Access, vol. 9, pp. 123672– 123688, Sept. 2021.

[15]. C. Dindorf, J. Konradi, C. Wolf, B. Taetz, G. Bleser, J. Huthwelker, F. Werthmann, E. Bartaguiz, J. Kniepert, P. Drees, U. Betz, and M. Fröhlich, “Classification and automated interpretation of spinal posture data using a pathology-independent classifier and explainable artificial intelligence (XAI),” Sensors, vol. 21, no. 18, p. 6323, Sept. 2021.

[16]. Z. Papanastasopoulos, R. K. Samala, H. P. Chan, L. Hadjiiski, C. Paramagul, M. A. Helvie, and C. H. Neal, “Explainable AI for medical imaging: Deep-learning CNN ensemble for classification of estrogen receptor status from breast MRI,” in Proc. SPIE 11314, Medical Imaging 2020: Computer-Aided Diagnosis, Houston, Texas, United States, pp. 113140.