Department of Computer Science and Engineering
(Data Science)

Privacy-Preserving Medical Analysis with Federated Learning and XAI

GUIDE : DR. N. SURESH KUMAR

TEAM - 4 :

ARJUN RAJESH MANVAR
(20BTRCD020)
PRAJWAL ARJUN SONKAVDE
(20BTRCD041)
PRATHAM AGARWAL
(20BTRCD042)

Content Outline

01. Objective

02. Abstract

03. Literature Review

04. Software and Hardware Requirements

05. Existing System

06. Proposed System

07. Methodology with Architecture Diagram

08. Architecture Diagram

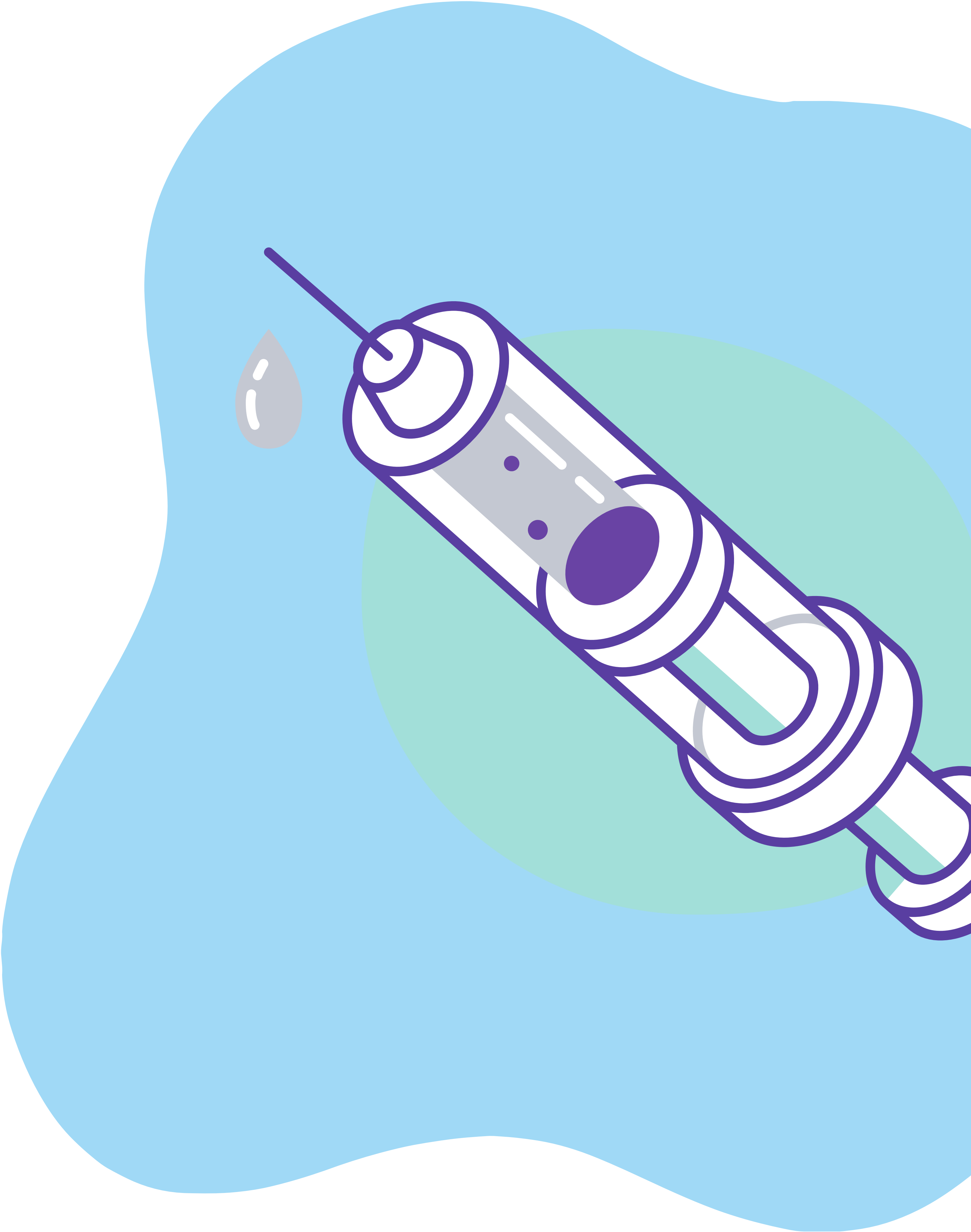
09. References

Objective

The objective of this study is to develop an Explainable AI model using federated learning in the healthcare domain. The proposed model aims to improve the accuracy and transparency of machine learning models while protecting the privacy of patients' data. Specifically, we will apply federated learning techniques to train machine learning models on decentralized healthcare data from multiple hospitals. We will also incorporate model-specific techniques for improved interpretability and transparency. The proposed model will be evaluated using real-world healthcare datasets to demonstrate its effectiveness in predicting patient outcomes while providing interpretable explanations of the model's decisions. This study will contribute to the development of Explainable AI models that can be used in healthcare settings while protecting the privacy and security of sensitive patient information.

Abstract

The aim of this project is to develop a machine learning system that can analyze medical reports from multiple decentralized sources while also providing transparent and interpretable results. The proposed system will use Federated Learning to train the model on data from multiple hospitals or clinics without centralizing the data, thus ensuring the privacy and security of patient data. To ensure interpretability of the model, the project will incorporate XAI techniques such as LIME and SHAP, enabling the model to provide explanations for its decisions. The project will evaluate the proposed system on a dataset of medical reports and compare its performance to other state-of-the-art models. The results of the evaluation will demonstrate the feasibility and effectiveness of using Federated Learning with XAI for medical report analysis, providing a valuable tool for medical professionals. The system can be used for various tasks such as predicting patient outcomes, identifying risk factors, and improving the accuracy of diagnoses.



Literature Review

Federated Learning for Healthcare Informatics" by Kairouz et al. (2019)

The authors present an overview of Federated Learning and its potential applications in healthcare informatics, including personalized medicine, predictive analytics, and clinical decision support systems. They also discuss the challenges of implementing Federated Learning in the healthcare domain, such as privacy concerns and data heterogeneity.

"Explainable Artificial Intelligence for Medical Image Analysis: A Review" by Zhou et al. (2021)

"Explainable AI for Medical Diagnosis: A Review" by Zhang et al. (2021)

The authors review various XAI techniques for medical diagnosis, including LIME, SHAP, and Integrated Gradients. They discuss the applications of XAI in medical diagnosis and the challenges of implementing XAI in the medical domain, such as the need for large annotated datasets.

"Federated Learning with Differential Privacy: Algorithms and Applications" by Li et al. (2020)

Software and Hardware Requirements

Software Components

1. Python
2. Scikit-learn
3. Matplotlib
4. Seaborn
5. Numpy
6. Pandas
7. Lime

Hardware components

1. High-performance CPUs or GPUs: The training process in federated learning requires high-performance CPUs or GPUs for efficient computation and faster processing of data.
2. Sufficient RAM: The system should have sufficient RAM to store and access the data during the training process.
3. Storage capacity: The system should have enough storage capacity to store the large amounts of data generated during the training process.
4. Network bandwidth: Federated learning involves the transfer of large amounts of data between devices, so a high-speed network connection is required to ensure smooth communication between devices.
5. Secure and reliable hardware: Federated learning involves the transfer of sensitive data, so the hardware used should be secure and reliable to prevent data breaches.
6. Clients (e.g., computers, mobile devices) to participate in the federated learning process

Existing System

The existing system of Explainable AI models can be broadly categorized into two types: model-specific and model-agnostic methods.

Model-specific methods are designed to work with specific types of machine learning models, such as decision trees or neural networks, and provide explanations that are tailored to the specific characteristics of the model. For example, a decision tree-based explanation might show the path of decisions that the model took to arrive at a particular prediction.

Model-agnostic methods, on the other hand, do not require knowledge of the specific model being used and can be applied to any machine learning model. These methods generally provide more general explanations of how the model arrived at its predictions, such as identifying which features were most important in making a decision.

Existing System

Some of the existing methods for Explainable AI include:

1. Local Interpretable Model-agnostic Explanations (LIME): a model-agnostic method that provides explanations for individual predictions by generating a local interpretable model around the prediction.

2. SHapley Additive exPlanations (SHAP): a model-agnostic method that uses game theory to identify the contribution of each feature to a particular prediction.

3. Integrated Gradients: a model-agnostic method that assigns importance scores to each feature based on how much it contributes to the model's prediction.

4. Decision tree-based explanations: model-specific methods that use decision trees to provide explanations of how the model arrived at its decision.

5. Attention mechanisms: model-specific methods that use attention mechanisms to identify which parts of an input were most important in making a prediction.

Existing System

Existing system for federated learning :

1. FATE: Federated AI Technology Enabler, is an open-source framework that supports federated learning on heterogeneous data sources. It has been used in various healthcare applications such as patient prediction models and medical imaging analysis.

2. Google's Federated Learning for Mobile Health : is a system that enables machine learning on mobile devices by aggregating model updates from multiple devices, while maintaining data privacy.

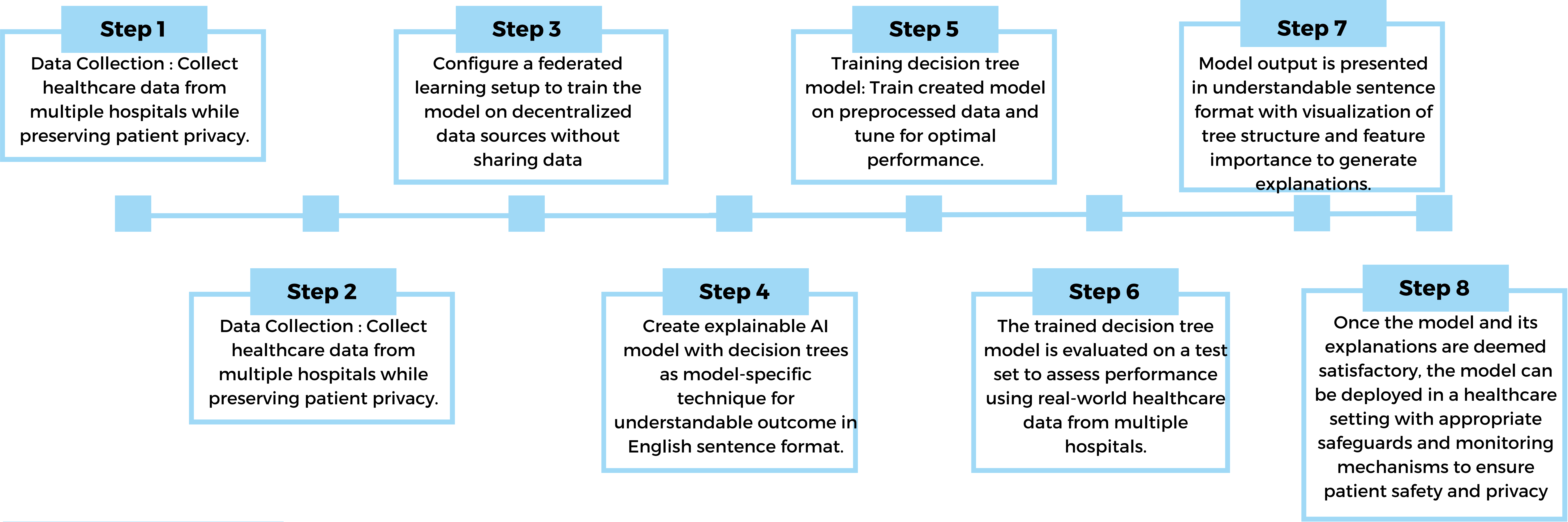
3. FEDERATEDHEALTH : is a federated learning system for electronic health record data analysis. It enables the collaborative learning of models across multiple institutions while ensuring privacy of patient data.

4. Federated Learning for Brain Tumor Segmentation: is a system that uses federated learning to train models for brain tumor segmentation on MRI data. It enables multiple institutions to contribute data while maintaining data privacy.

Proposed System

- Medical report analysis is critical for informed decision-making in healthcare.
- Machine learning has shown promise for medical report analysis.
- Centralized storage of medical data raises privacy and security concerns.
- Federated Learning is a solution that allows decentralized training of machine learning models while preserving data privacy.
- Explainable AI (XAI) techniques provide interpretability to the model's decisions, making it easier for medical professionals to understand and trust the model's predictions.
- A proposed system will use Federated Learning with XAI techniques to develop a machine learning system for medical report analysis.
- Federated Learning with XAI techniques can improve the accuracy, generalizability, and interpretability of the model while preserving patient privacy and data security.
- The proposed system has the potential to be used in various medical settings, leading to better patient outcomes through improved diagnoses and treatments.

Methodology



Step 1

Data Collection : Collect healthcare data from multiple hospitals while preserving patient privacy.

Step 2

Data Collection : Collect healthcare data from multiple hospitals while preserving patient privacy.

Step 3

Configure a federated learning setup to train the model on decentralized data sources without sharing data

Step 4

Create explainable AI model with decision trees as model-specific technique for understandable outcome in English sentence format.

Step 5

Training decision tree model: Train created model on preprocessed data and tune for optimal performance.

Step 6

The trained decision tree model is evaluated on a test set to assess performance using real-world healthcare data from multiple hospitals.

Step 7

Model output is presented in understandable sentence format with visualization of tree structure and feature importance to generate explanations.

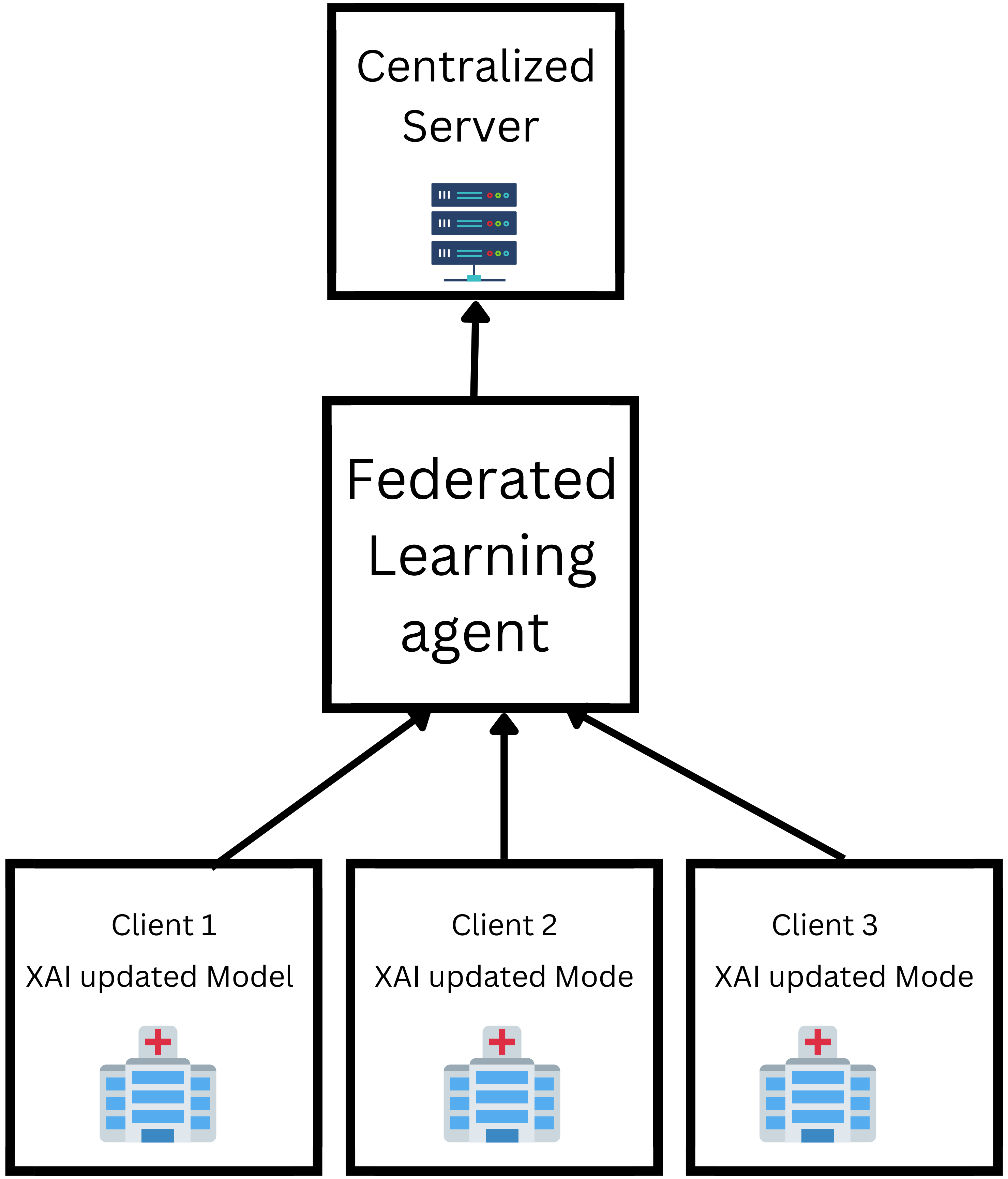
Step 8

Once the model and its explanations are deemed satisfactory, the model can be deployed in a healthcare setting with appropriate safeguards and monitoring mechanisms to ensure patient safety and privacy

Expected Outcomess :

1. Development of an explainable AI model using a model-specific method and federated learning in healthcare domain.
2. Demonstration of the efficacy of the proposed system in predicting patient outcomes while preserving patient privacy.
3. Potential applications in personalized medicine and population health management.

Architecture Diagram



References

- Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... Zhang, Z. (2019). Advances and Open Problems in Federated Learning. arXiv preprint arXiv:1912.04977.
- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning Deep Features for Discriminative Localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 2921-2929).
- Zhang, Y., Chen, P., Liu, B., Kermany, D. S., & Chen, H. (2018). Towards Explainable Deep Learning for Diagnosis of Acute Myeloid Leukemia. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD) (pp. 2596-2604).
- Li, S., Wang, S., Xu, K., & Zhao, J. (2020). Federated Learning with Differential Privacy: Algorithms and Applications. IEEE Communications Magazine, 58(10), 64-69. doi: 10.1109/MCOM.001.2000175.
- Kuo, T.-T. and Pham, A. (2022) 'Detecting model misconducts in Decentralized Healthcare Federated Learning', International Journal of Medical Informatics, 158, p. 104658. doi: 10.1016/j.ijmedinf.2021.104658.