Milestone 2 - Progress Evaluation

1) Project Title: Tomographic Medical Image Reconstruction using Federated Learning
   Members: Joshua Sheldon (jsheldon2022@my.fit.edu), Yash Jani (vjani2023@my.fit.edu), Tanuj Kancharla tkancharla2022@my.fit.edu), and Izzy MacDonald (imacdonald2022@my.fit.edu)
2) Faculty Advisor: Dr. Debasis Mitra (dmitra@fit.edu)
3) Client: Same as advisor.
4) Progress matrix for current milestone:

Task	Completion	Joshua	Izzy	Tanuj	Yash	To Do
Orchestrator application user interface	80%	10%	40%	10%	40%	Need to connect the frontend to the backend
Orchestrator application initial model selection	50%	40%	10%	10%	40%	Design and API route done, just need to implement
Contributor Application can accept training data	50%	25%	25%	25%	25%	Design and API route done, just need to implement
Correct and accurate synthetic data can be generated	100%	25%	25%	25%	25%	None

5) Task Discussions
   a) Learning managers needed a way to interact with the model, and a way to train the model. We created a user-friendly interface that allows the learning manager to see the different models that have been trained, having the choice to remove them as well. The user can also choose an initial model to train, choosing to upload data or with random parameters. They can see which contributors are available, see their sample counts and when they most recently uploaded. This information allows them to choose which contributors to train the model on, and

then commence the round. This is only a frontend model at the moment, we have not connected the backend.

b) Since training may not be done on the model initially, the model needs to be able to randomly initialize one. We have a frontend for this task, however, not a backend yet. However, the application now has a class diagram, and we have a clear idea of how we'll implement this functionality.

c) The contributor applications need a mechanism for users to select which data will be used for training. We have an API route for this task in the application, but the functionality has not been implemented. However, the application now has a class diagram, and we have a clear idea of how we'll implement this functionality.

d) We want to ensure that the pipeline is generating correct and accurate synthetic data, along with having some variety. We are looking into some variation with our synthetic data through introducing lesions in the heart, along with having more heart/organ models to generate data with. Izzy and Joshua were able to look into past data and were able to implement lesions into the code. Allowing for heart generations to include lesions. Joshua was able to reach out to Dr. Paul Seggars and get more base models to generate data with. We were able to implement it into the code to include a randomization of patient heart/organ files along with the parameter randomization.

6) Member Contributions

Joshua:
- Used SSIM to get the average similarity between any two sinograms from the same XCAT phantom.
- Created UI mockups for the orchestrator and contributor applications.
- Revised system architecture diagram.
- Made class diagrams for the orchestrator and contributor applications.
- Started work on a data synthesis pipeline class diagram.
- Made sequence diagram learning for our federated learning implementation.
- Deciphered the current data synthesis pipeline and made a refactoring plan.
- Expanded the design document from 9 to 19 pages.
- Created initial versions of the orchestrator and contributor applications.

Izzy:
- Created the orchestrator UI
- Investigated lesion introduction to XCAT phantom generation
- Created XCAT phantom and XCAT/XCAT+ generations with lesions
- Helped implement parameter randomization into the code when generating XCAT phantoms

- Implemented randomization of patient heart/organ data into the code when generating XCAT phantoms

Tanuj kancharla :

- Investigated free, secure storage options that support user authentication, encrypted communication, local downloads, and incremental updates.
- Developed a Prototype for a Storage System in code
- Evaluated Data Storage Options for Federated Learning
- Researched different encryptions and authentication solutions for optimal data transfer and security

Yash:

- Refactored the augmentation script to make it a lot cleaner and removed unnecessary transpositions of data
- Optimized the augmentation script for memory usage and changed/removed transformations on augmentation to increase the quality of artificial data
- Implemented parallelization, specifically multiprocessing, to make the augmentation process N times faster, dependent on the # of CPU cores
- Changed and optimized a couple of scripts on the initialization process of the simulation to allow it to pick up where it stops in case of a crash or failure
- Investigating how to squash OpenGATE's returned dimensional output from 128x120x240 -> 128x120x120

7) Next Milestone Plan

Task	Joshua	Izzy	Tanuj	Yash
Orchestrator application UI	10%	70%	10%	10%
Orchestrator application initial model selection	40%	10%	10%	40%
Contributor Application can accept training data	25%	25%	25%	25%
Implement new Orchestrator Features	10%	40%	10%	40%

8) Task Discussions
    a) This is a rollover task from the last milestone, the frontend is up and running. However, it does not connect to real data. The goal for the next milestone is to connect it to the backend.
    b) This is a rollover task from the last milestone. With this milestone the objective is to get a backend up and running for the orchestrator application. With a class diagram, we have a clear understanding of how to implement this.
    c) This is a rollover task from the last milestone. The objective of this milestone is to get the contributor application up and running, so that the contributor can upload data. Once again, the main focus is simply that training data can be uploaded, not that the model can then be trained on the uploaded data.
    d) Now that the Orchestrator application has a user interface, there are certain features we want the application to be able to do. The learning managers should be able to define which contributors they deem trustworthy, see which contributors have new data for training, along with starting a round of training and selecting which contributors to participate in it. These features will be able to be implemented once we have connected the backend to the frontend.
9) Meeting Dates with Client: See Meeting Dates with Advisor.
10) Client Feedback: See Faculty Advisor Feedback below.
11) Meeting Dates with Advisor: 3/7 9am-10am, 3/14 1pm-2pm
12) Faculty Advisor Feedback: Reduce bureaucratic overload (too much details and documents burden the primary task). Focus on the broader picture of image reconstruction and the challenges therein.

Faculty Advisor Signature: _____ Date: _____

13) Evaluation by Faculty Advisor

- Faculty Advisor: detach and return this page to Dr. Chan (HC 209) or email the scores to pkc@cs.fit.edu
- Score (0-10) for each member: circle a score (or circle two adjacent scores for .25 or write down a real number between 0 and 10)

Joshua Sheldon	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Yash Jani	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Tanuj Kancharla	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Izzy MacDonald	0	1	2	3	4	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10

Faculty Advisor Signature: _____ Date: _____