Client Meeting

Date: September 22nd, 2021
Attendees: Rasika, Quinn, Emily, Niklas, Kyle, Alex, Joel, Marieke, Josh, Zach
Absent:

Agenda:

- Overview of Project
  - Curating a large database of mammograms from various public and private facilities throughout the province
  - Mismatches are non-identical matches of incoming studies with existing study data
  - Simple example:

Name: DOE^JANE DOB: 19501515 PatientID: 1337ID	Name: DOE^JAYNE DOB: 19501515 PatientID: 1337ID

    -
    - flag as potentially mismatched images.
  - In house mismatch codes:

2 – Most info matches, different DOB and first name
3 – Most info matches, different DOB and last name
4 – Most info matches, different DOB
5 – Most info matches, different patient ID
6 – Most info matches, different patient ID and first name
7 – Most info matches, different patient ID and middle name(s)

    -
  - There are scenarios where these can be very difficult to resolve with confidence.
    - Compression complicated this issue due to the skin possibly being compressed in a slightly different way when the images are taken.
    - Images can come from either screenings or diagnostics so there may be missing images. ie. diagnostic might only have one image.
  - We want a model that will take two mammograms and give us the probability of it being from the same patient.
  - cleaning and cataloging the dataset is what was meant.
- Features:
  - What features are required?
  - What tech stack would you prefer?
    - Python
    - Rest is basically up to use to research - Tensorflow?
- Acceptance Criteria:
  - What is the general acceptance criteria of the project?
    - Training dataset and testing set with a good performance level on the test dataset for the model. (70-90%)?
    - If we can get meaningful results, they're happy
- General Questions:

- Where is the data currently hosted?
- When will we get access to the data?
    - In the beginning we will be given a different dataset and will be given the medical data eventually (after the Health Canada Course)
- Any restrictions as to when and how we work?
    - What days do we have to be there?
        - We have to be there at least once to activate our accounts.
        - Then we will have remote access with Citrix
    - How many hours a week are expected?
- When is the Health Canada Course on Research Ethics -TCPS 2: CORE (https://www.tcps2core.ca/welcome) due?
    - early stage, complete and send to Rasika before we are able to use real data
- Would you like a project proposal with a detailed tech stack and plan written up? If so, when would you like this to be done by?
- For our git repos, what privacy settings/licensing should we apply?
    - so long as we exclude any data they don't care about code.
    - summary data and extracts are not considered protected data.
    - images cannot leave BC Cancer

- Action Items:
    - **Send full names to Rasika/Quinn > Client Liaison** by ?
    - **Write up and send Client Meeting Agenda to Quinn/Rasika** > Client Liaison by Monday
        - Add Questions/Notes to Agenda > All by Monday
    - Define features
    - Git Repo set up
    - Milestone layout (which features when) done in week 4.
    - Develop testing strategy
    - Identify tech stack