Diagnosing Lipohypertrophy at the Bedside

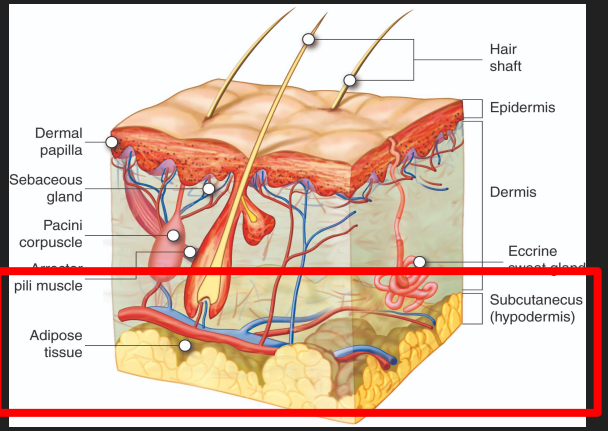
Students – Ela Bandari, Javairia Raza, Lara Habashy and Peter Yang

Mentor – Tomas Beuzen

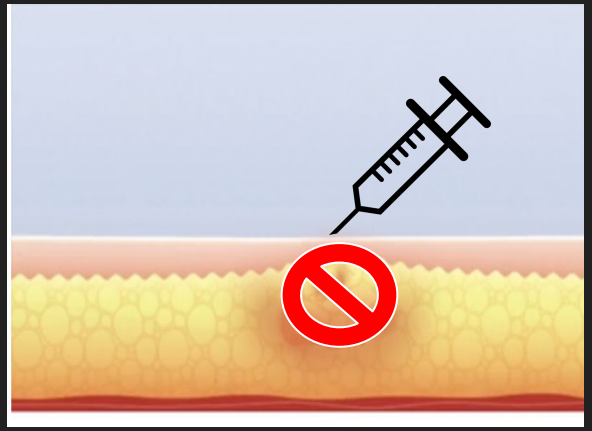
Capstone Partner – Dr. Ken Madden, Gerontology and Diabetes Research Laboratory (GDRL)

What is lipohypertrophy?

- ***Lipohypertrophy*** is a common complication for diabetic patients who inject insulin



Occurs in the deepest layer of the skin



Classified as the development of fat cells, fibrous tissue and decreased veins at injection site

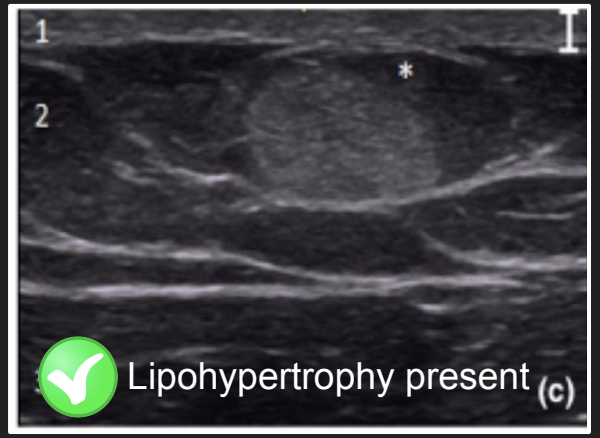
Patients are unable to manage blood sugar levels

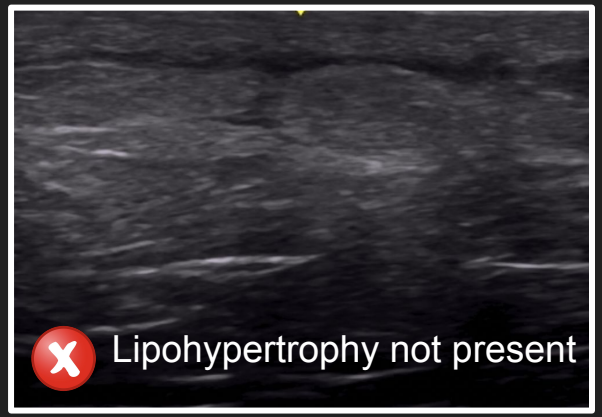
May need more insulin

Big Question

- Research has suggested that ***ultrasound imaging*** is much more effective than physical examination of the body
- The criteria used to classify is currently implemented by a small group of physicians only

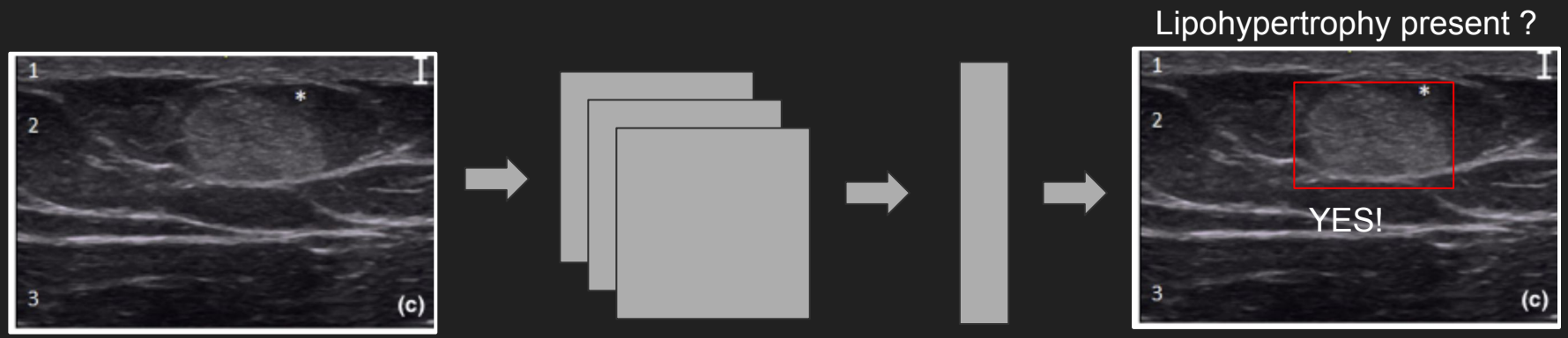
Can we leverage supervised machine learning techniques to accurately classify the existence of lipohypertrophy given an ultrasound image?





Refining the Question for Data Science

- Can we develop a ***binary classification CNN model*** that will accurately classify ultrasound images?
- Can we leverage ***object detection*** techniques to classify, given a positive lipohypertrophy site, the exact area of its location on a given ultrasound image?



Target Audience

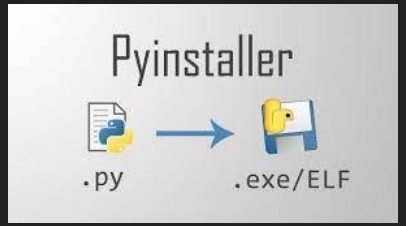
- **Primary Audience:** Healthcare professionals
  - Unfamiliar with Programming
- **Secondary Audience:** Researchers and Ultrasound Manufacturers
  - Proof of Concept

Data Product

1. Source code to be integrated into the device

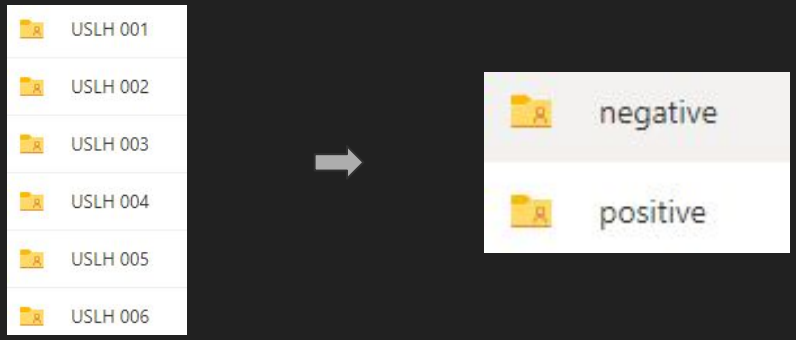
2. An executable script hosted on the cloud

3. An executable script that can be run locally



Data Wrangling

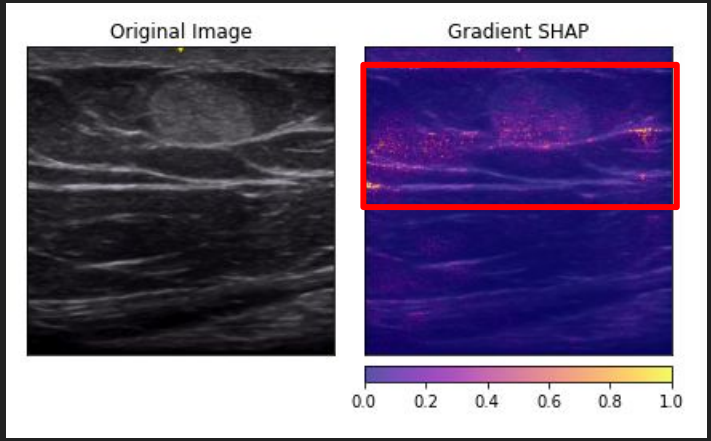
- Classify and rearrange images



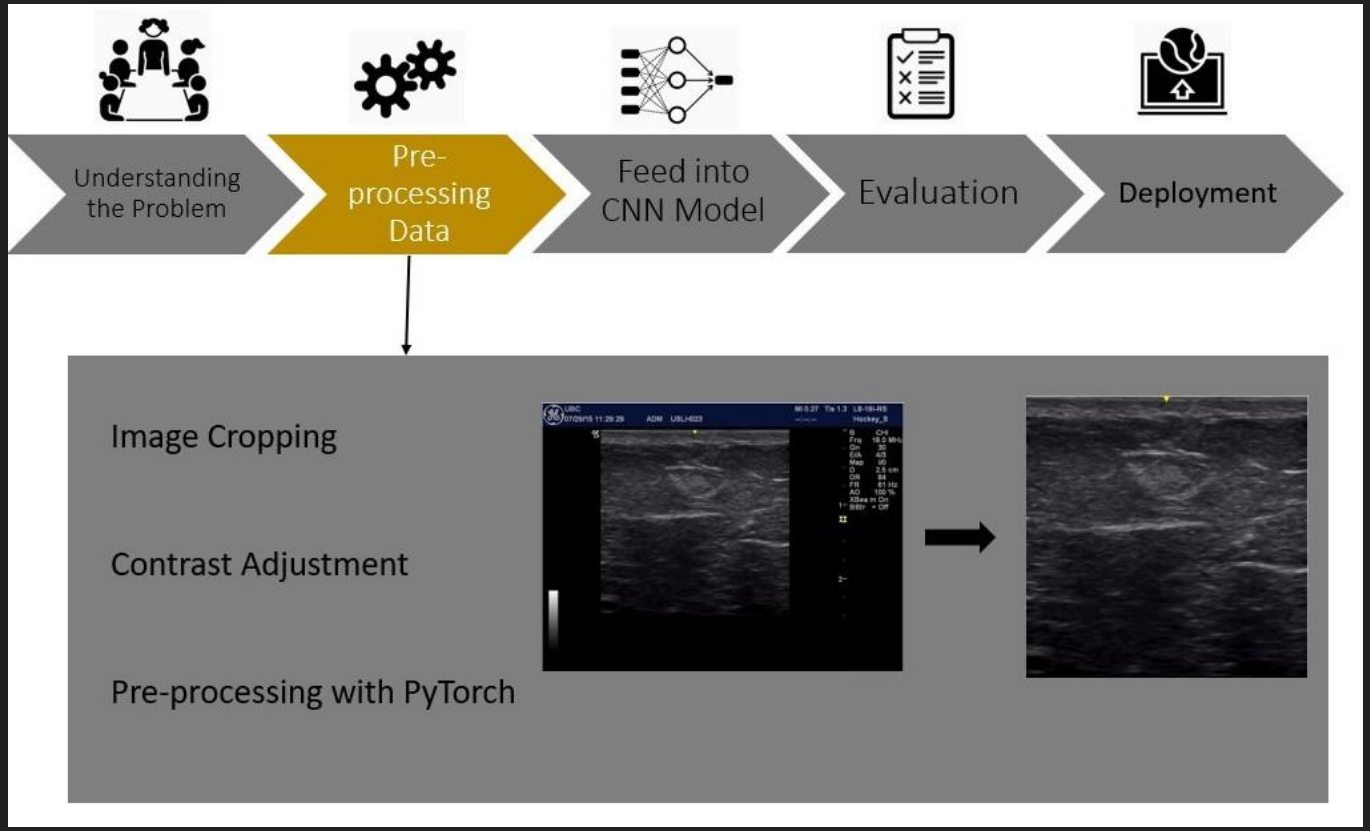
- Crop images

EDA and Preliminary Findings

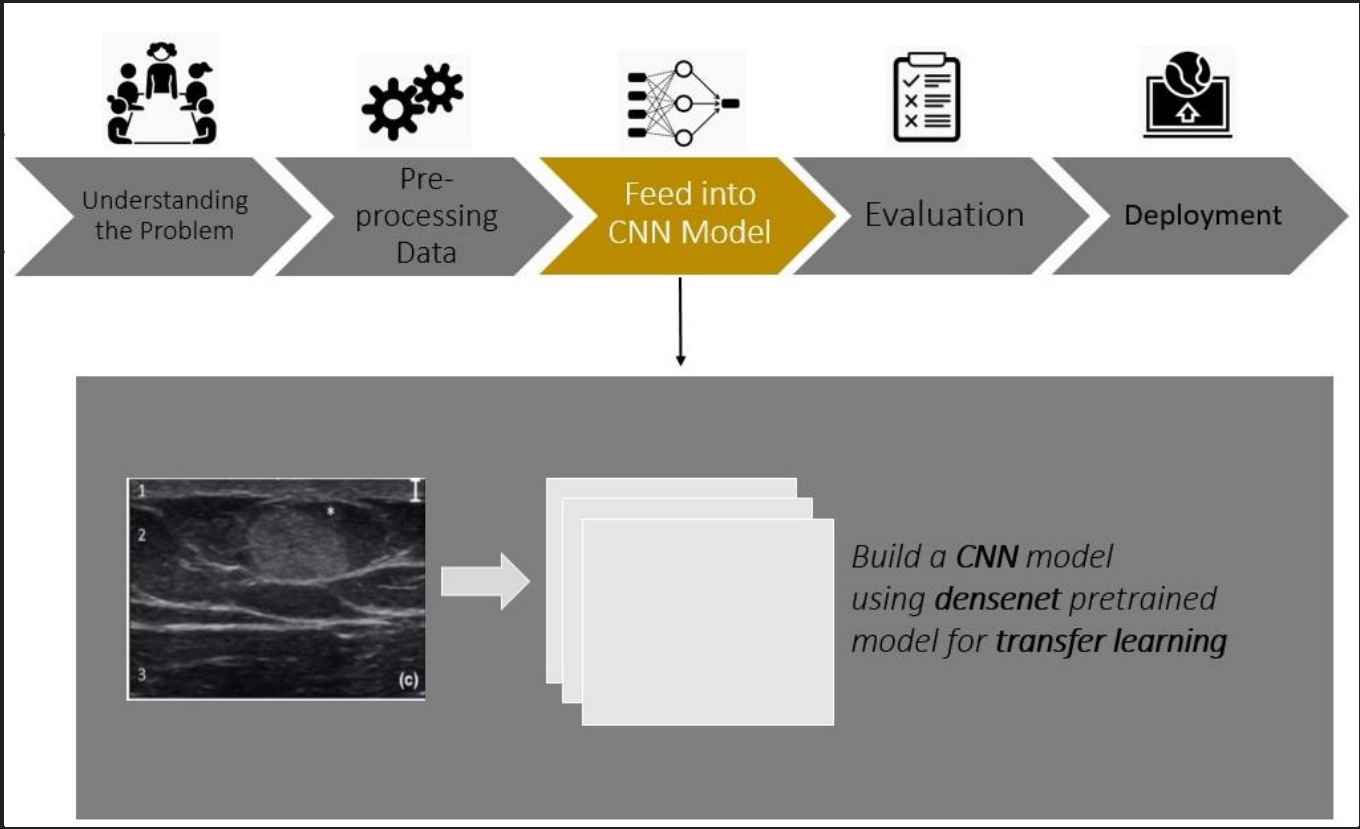
- ***263*** Ultrasound images
- Balanced target, with ***51%*** positives and ***49%*** negatives
- ***82%*** validation accuracy score from densenet after 20 epochs
- Feature importance of the trained densenet makes sense



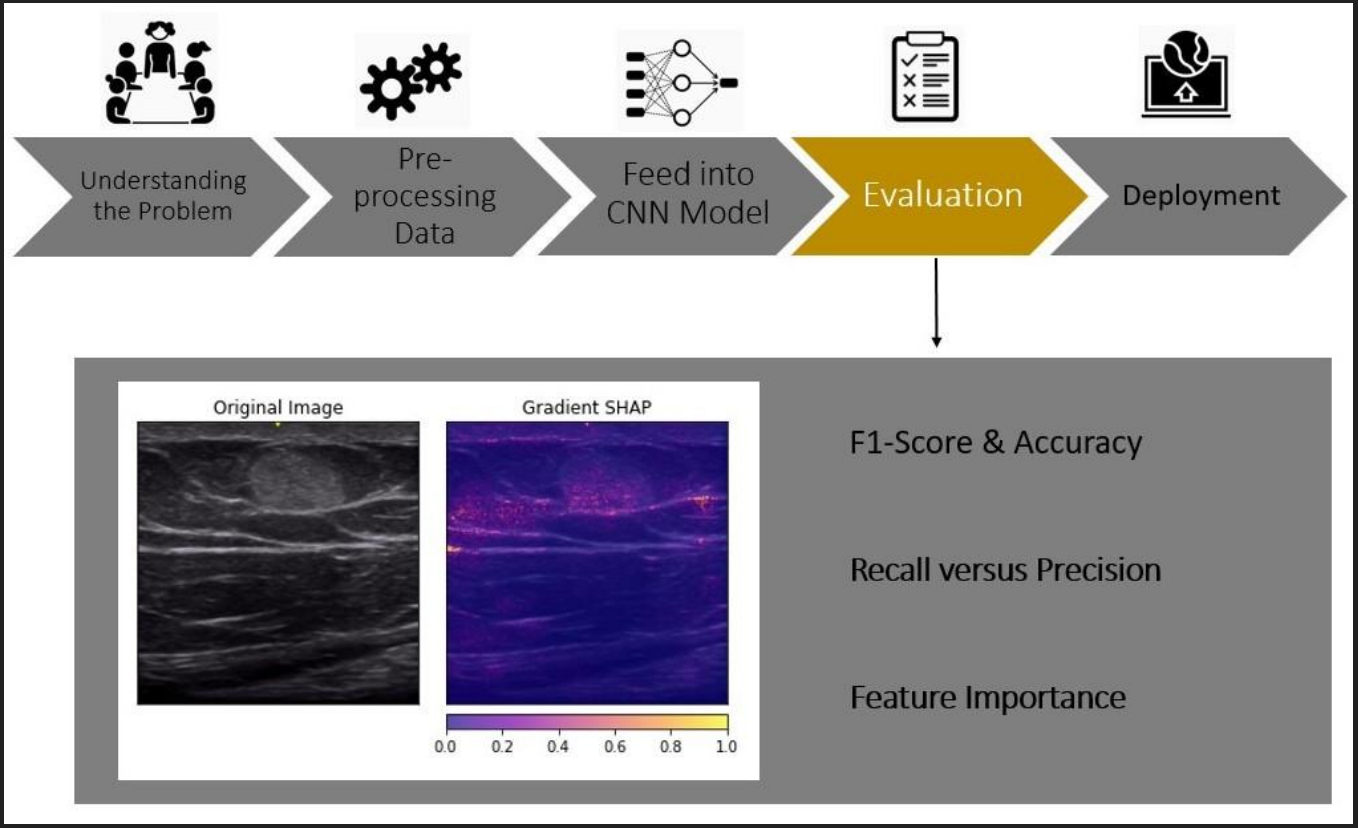
Data Science Techniques: Pipeline



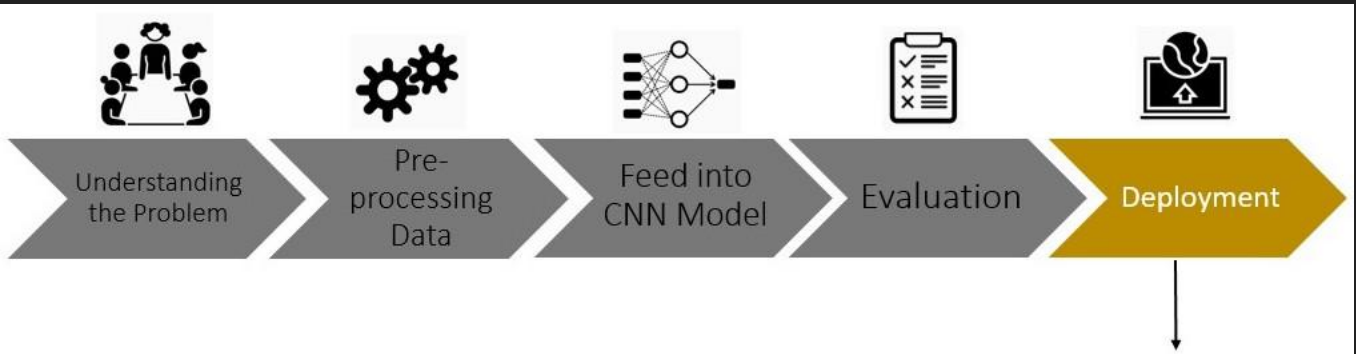
Data Science Techniques: Pipeline



Data Science Techniques: Pipeline



Data Science Techniques: Pipeline



Obtain prediction result as

1. **Lipohypertrophy** detected

OR

2. **Lipohypertrophy** not detected

Data Science Techniques: Exploring Other Architectures

- Baseline Model: CNN using the pretrained model **densemodels**
  - **Problem: Requires high GPU memory + long training time**

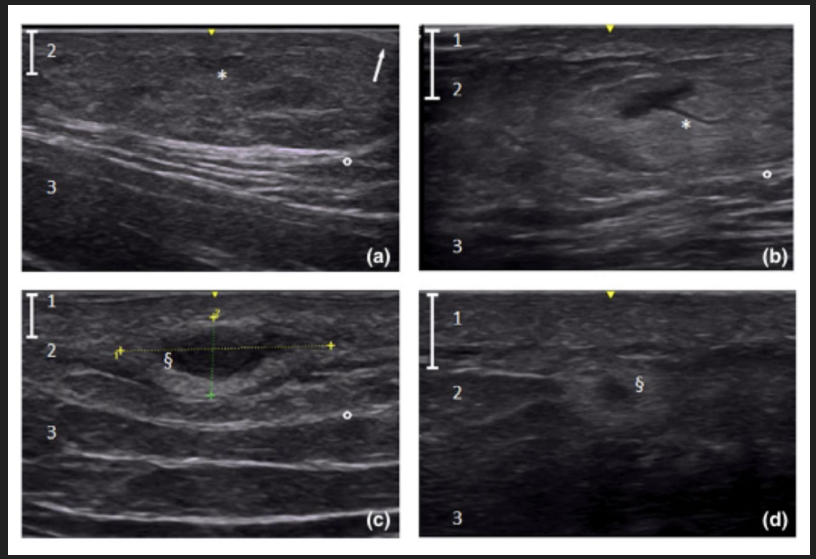
- Alternative Architectures
  - ***VGG*** proposed in the paper *"Very Deep Convolutional Networks for Large-Scale Image Recognition" (2014)*
    - Pro: Reduced the number of parameters, very effective at extracting features from images
    - Con: Slow to train, learned weights are quite large

  - ***Inception*** proposed in the paper *"Going Deeper with Convolutions" (2015)*
    - Pro: smaller weights, better handles images with more depth

  - Many other architectures and variants of mentioned architectures!

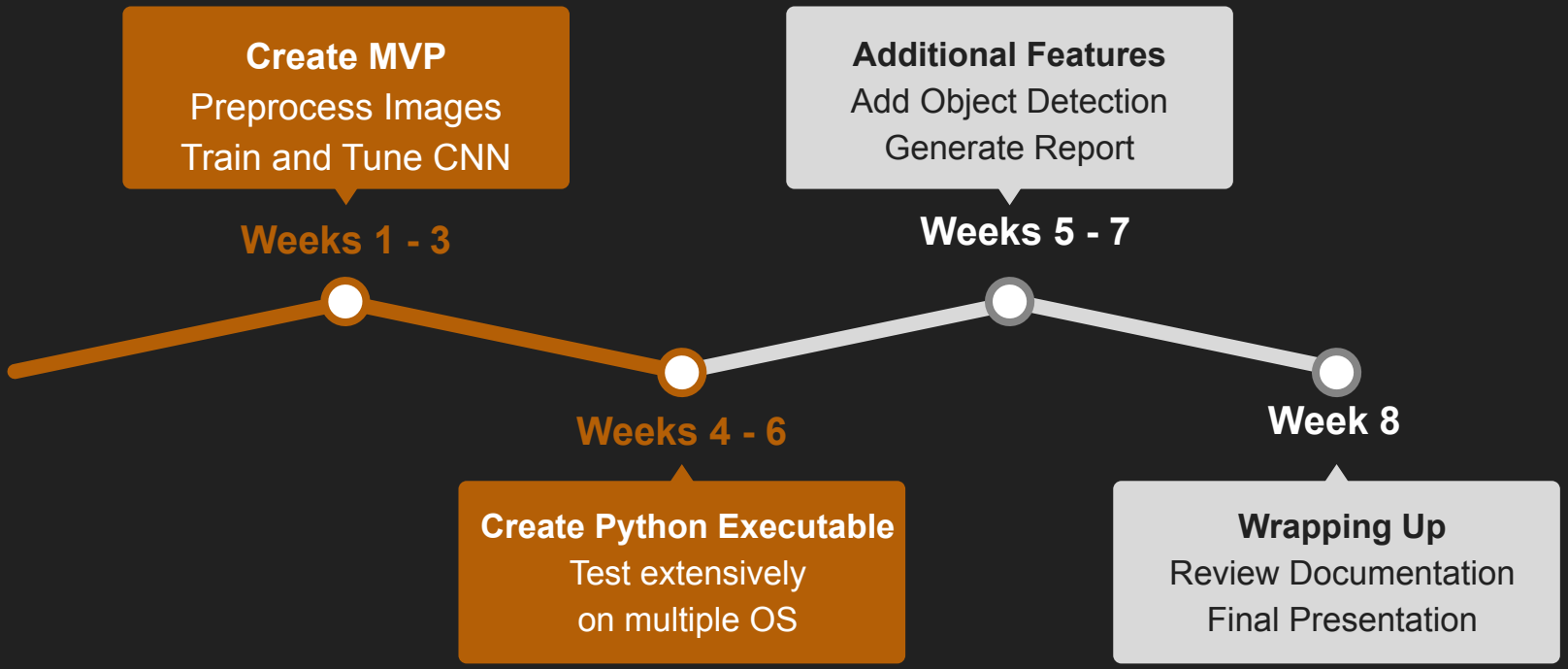
Anticipated Problems

An abdomen area where insulin is ***never injected***



Abdomen areas without lipohypertrophy at sites ***injected*** with insulin

Timeline


Create MVP
Preprocess Images
Train and Tune CNN
Weeks 1 - 3
Additional Features
Add Object Detection
Generate Report
Weeks 5 - 7
Weeks 4 - 6
Create Python Executable
Test extensively
on multiple OS
Week 8
Wrapping Up
Review Documentation
Final Presentation

Thank you for listening!
Any questions?