November 29 - Project Kickoff

So today was the official start of my lung segmentation project! I went hunting for datasets and found a great one on Mendeley that's got a ton of chest X-rays. Here's the breakdown:

- **Darwin Dataset**: 6,106 images (includes viral pneumonia, tuberculosis, and normal lungs)
- **Montgomery Dataset**: 139 images
- **Shenzhen Dataset**: 566 images

These images will help me identify different lung conditions, which is exactly what I need for segmentation.

I also found a **mask image** for lung segmentation. This is going to be super useful because my goal is to detect and segment the damaged parts of the lungs.

Starting Research

I wanted to make sure I understood the diseases I'd be working with, so I took some time to do a deep dive into **viral pneumonia** and **tuberculosis**. Here's what I learned:

- **Viral Pneumonia**: This is an infection caused by several viruses like influenza and even coronaviruses (hello, COVID-19!). The symptoms include fever, cough, and fatigue. On X-rays, the lungs look cloudy, which is something I need to watch for in my images.
- **Tuberculosis (TB)**: TB is a bacterial infection that affects the lungs and is highly contagious. It causes symptoms like a persistent cough, chest pain, and night

sweats. On chest X-rays, it usually shows up as patchy lesions, which I'll need to identify.

- **Pneumothorax**: This is when the lung collapses because air gets trapped between the lung and chest wall. It can cause sharp chest pain and trouble breathing. On X-rays, you'll see a clear space around the lung. This is what I'm hoping to spot in my images.

November 30 - Reaching Out for Help

I decided to get some advice from the experts, so I emailed a few professors at the University of Calgary. I wanted to see if they could help guide me with some insights and maybe point me toward helpful resources. Here's the email I sent:

Subject: Request for Assistance with Lung Segmentation Project

"Hi Dr. _____,

My name is Anika Rastogi, and I'm a grade 10 student at Westmount Charter School. I'm working on a science fair project focused on lung segmentation, and I'm trying to develop an AI to scan lungs for damaged tissue."

"I came across your work and believe your expertise could be really helpful for my project. If you have some time, I'd love to get your advice or any resources you can share!"

"Thanks so much for considering my request! I'm looking forward to hearing from you!"

Best,

Anika Rastogi

On the coding side, I got the **preprocessing** part done! I loaded up the images and masks, resized everything to 256x256 pixels, and normalized the pixel values. Feels good to be moving forward!

December 1 - Responses Rolling In

I was so excited to see replies from a bunch of professors! They shared some really helpful advice and pointed me to some great articles on **AI in medical imaging** and **deep learning**. Definitely going to spend some time reading through them. But they did say that my code seemed pretty accurate.

One professor response;

"Anika,

I was on call all last week so wasn't able to look at emails.

Yes, the black part is the damaged tissue. It's tough though because what it outside the area of the lung is also black so it's hard to tell the edge of the lung (and therefore where the damage stops)."

December 27-30 - Moving On to Visualization

I wanted to see how my model was performing, so I created **accuracy** and **loss plots** to track the progress. It's really satisfying to see those curves moving in the right direction! The plots are giving me a clear picture of how well my model is learning, so I'm excited to keep going.

January 1 - New Year, New Dataset!

I came across an **extra dataset** with **pneumothorax images** from Kaggle from a competiton! It also included preprocessed masks for segmentation, which is exactly what I needed. I was pumped about this because it's such a perfect fit for my project.

January 2 - Debugging Problems

Today, I ran into some issues with **VS Code** while trying to load the pneumothorax dataset. The app kept crashing! After some troubleshooting, I realized that the problem was the **huge file sizes**. Once I removed the annotation files, everything worked just fine. Sometimes it's the small things!

January 15 - Expanding the Model

By now, I felt like my model had a good foundation, so it was time to **expand it**! I added more **filters** to the **CNN** and brought in an **encoder-decoder architecture** to improve the segmentation. This setup helps the model compress the image's features and then reconstruct them, which is perfect for what I need to do with lung segmentation.

January 16 - A Little Refiner Trouble

I tried adding a **refiner** to the model to make the segmented images sharper. Unfortunately, it didn't go as planned. The results were a little too blurry, and I wasn't

happy with how it turned out. For now, I've decided to leave it out and come back to it later.

January 18 - Research and Documentation

With most of the coding done, I turned my attention to **documentation**. I started writing notes to clarify the technical side of my code and the diseases I'm studying. It's also a good way to stay organized for the final report.

Learning About Deep Learning

I've been diving deep into **U-Net**, a popular architecture for **medical image segmentation**. It's got this cool encoder-decoder structure that works wonders for extracting and reconstructing features in images. I'm excited to keep using it in my project because it's perfect for lung segmentation tasks!

February 1 - Severity Levels Integration

Adding Severity Levels to the Model

Today, I decided to make things more interesting by adding **severity levels** to the diseases I'm working with. It felt like a natural next step because, instead of just detecting whether the lungs have damage, I could also classify how **severe** that damage is. This helps doctors know exactly how bad the infection is and what action to take.

For **pneumonia** and **tuberculosis**, I decided on three severity levels:

- **Mild**: Just a little bit of lung involvement, like in the early stages of infection.

- **Moderate**: More than **50%** of the lung is affected, showing it's a bigger problem.
- **Severe**: A lot of the lung is damaged—think major infection that needs urgent attention.

For **pneumothorax** (when the lung collapses), the severity is:

- **Mild**: Tiny air pockets with no big collapse.
- **Moderate**: Partially collapsed lung, still requires attention.
- **Severe**: Full collapse—this is the serious stuff, a medical emergency!

I started thinking about how I can measure features like **lung size** and **damage percentage** to make the classification work.

February 10 - Model Testing and Evaluation

Testing the Model

After adding the severity levels, I tested how well the model could tell the difference between **mild**, **moderate**, and **severe** conditions. The good news is the model did well with **mild**,**moderate, and severe** cases. Just to make sure my code was giving me proper answer, I did email professors again, asking for input.

February 15 - Preparing for Presentation

With the model doing better, I started thinking about the **science fair**. I needed to get my **presentation** ready, and I wanted to make sure it was not only informative but also fun to watch. So, I put together my slides and broke everything down:

- **Intro to Lung Diseases**: I explained why diseases like pneumonia, tuberculosis, and pneumothorax are such a big deal.
- **AI to the Rescue**: How AI can help doctors spot these diseases early by analyzing X-rays of the lungs.
- **Data & Preprocessing**: Talked about where my data came from and how I cleaned it up.
- **CNN Model**: I gave a simple explanation of how Convolutional Neural Networks (CNN) work and why they're good for this task.
- **Severity Classification**: How I split the diseases into severity levels and why that's important.
- **Results**: I showed the model's performance in terms of **accuracy** and **segmentation results**.
- **Future Work**: What could be improved and where I want to take this project next.

On top of this I did get an email back from the professor saying that the image I sent him was a severe image, which is exactly what my code said

March 1 - Final Model Review

Final Test Run

It was time for the final test. I ran the model through its paces one last time, ensuring it could correctly classify severity levels for all the conditions. The results were great—**accurate** and **consistent** across both the training and validation sets. I felt confident that it was finally ready.

March 5 - Presentation Practice

I started rehearsing for my **presentation**. I wanted to make sure everyone could understand how cool (and important) this project is, even if they don't know much about AI or medicine. I was focusing on being clear about how **AI** can help doctors detect lung diseases early, which is really the heart of my project.

March 15 - Final Touches

The website closes on March 21, and I still have lots to do. Logbook, 10 minute video and presentation (addressing the severity level of things I completed). During my coding class, I finished logbook edits, and adding 2 more slide to the slide show. But I still have the 10 minute video left which I left to do either today or tomorrow. But we will see.

March 17 - video

The video is finally finished and noW I just have to put all of it on the cysf website and wait for the city wide science fair!!

*** not able to add photos, the pictures are not formatted correctly to be put on google docs.