# FDA Submission

**Your Name:** Louise O'Connor

**Name of your Device:**

## Algorithm Description

### 1. General Information

**Intended Use Statement:** Predicting the presence of pneumonia given a chest xray to assist the radiologist.

**Indications for Use:** This algorithm was trained on male and female patients ages spanning 1 to 95 years who have beeen administered a chest x-ray. All patients were scanned in either Posterior-Anterior or Anterior Posterior positions.

**Device Limitations:** The presence of comorbidities in the chest xrays is a limitation of the algorithm. The presence of other diseases could affect the algorithms sensitivity and specificity and reduce the ability to accurately detect pneumonia. Some diseases are similar as pneumonia in terms of pixel distribution so only using pixel distribution is an algorithm limitation.

GPU and Cloud infrastructure would also be required for the device to achieve fast performance so this is a computational limitiation of the device.

**Clinical Impact of Performance:** False Positives would be detected by a clinician on a second pass. False negatives are more serious as it would lead to the missed diagnosis of pneumonia. As a result, it would be important to optimize recall.

### 2. Algorithm Design and Function

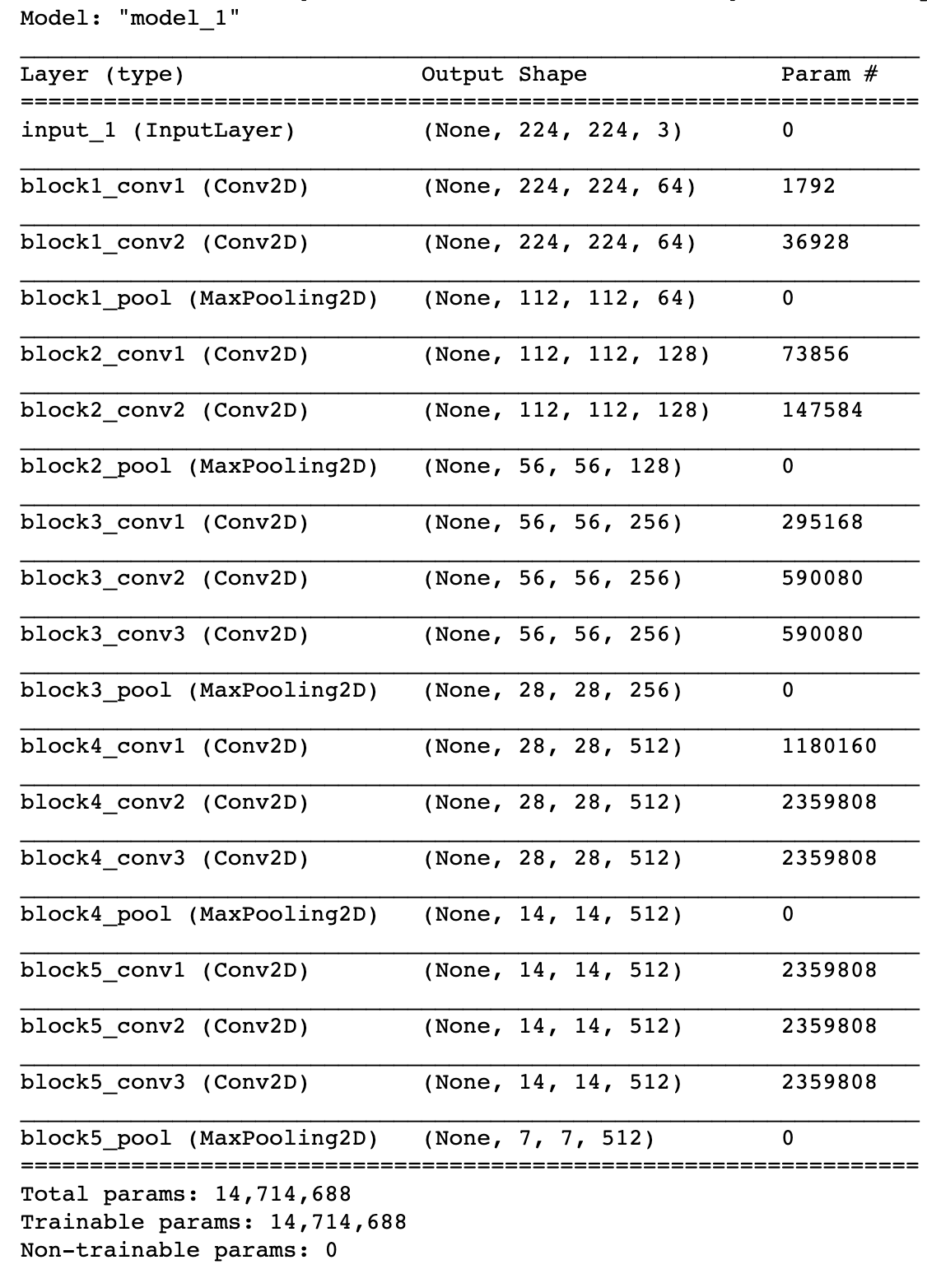
**DICOM Checking Steps:**

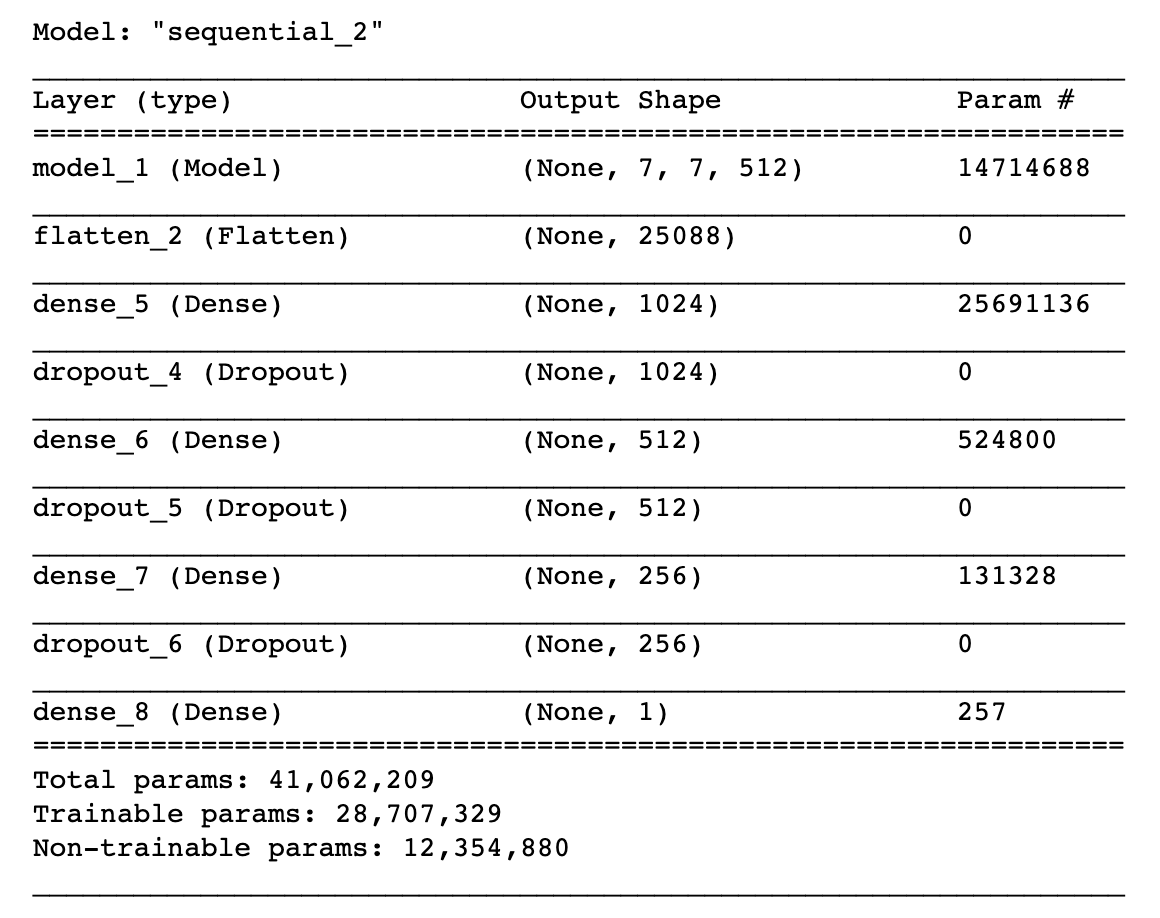
Reads in a .dcm file, checks the important fields for our device, and returns a numpy array of just the imaging data. Ensures that the image is was taken using the correct modality and the body part examined is the chest

**Preprocessing Steps:**

Takes the numpy array output by check\_dicom and rescales the image. Normalizes the image.

**CNN Architecture:**





### 3. Algorithm Training

**Parameters:**

Types of augmentation used during training:

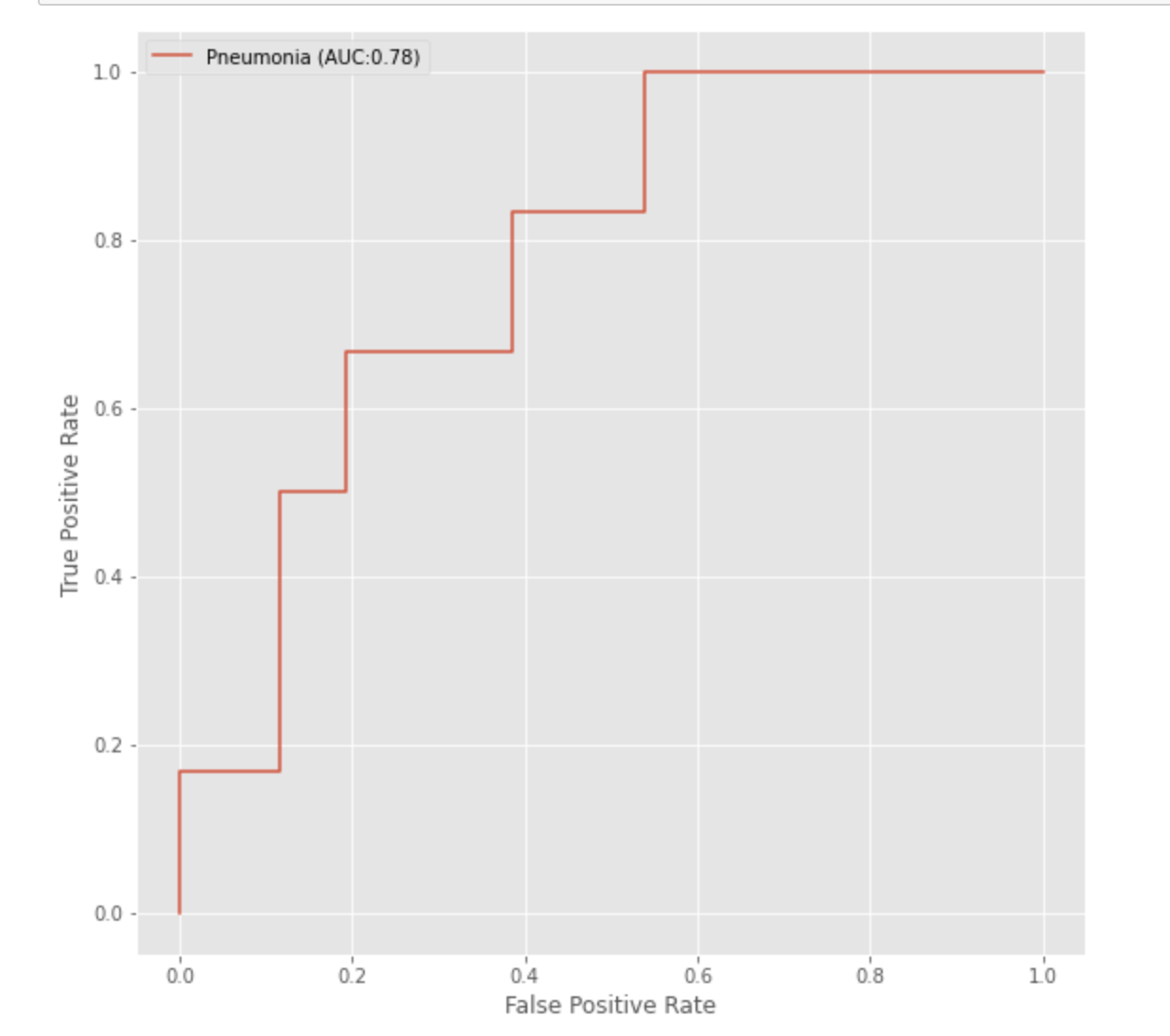
Batch size: 32

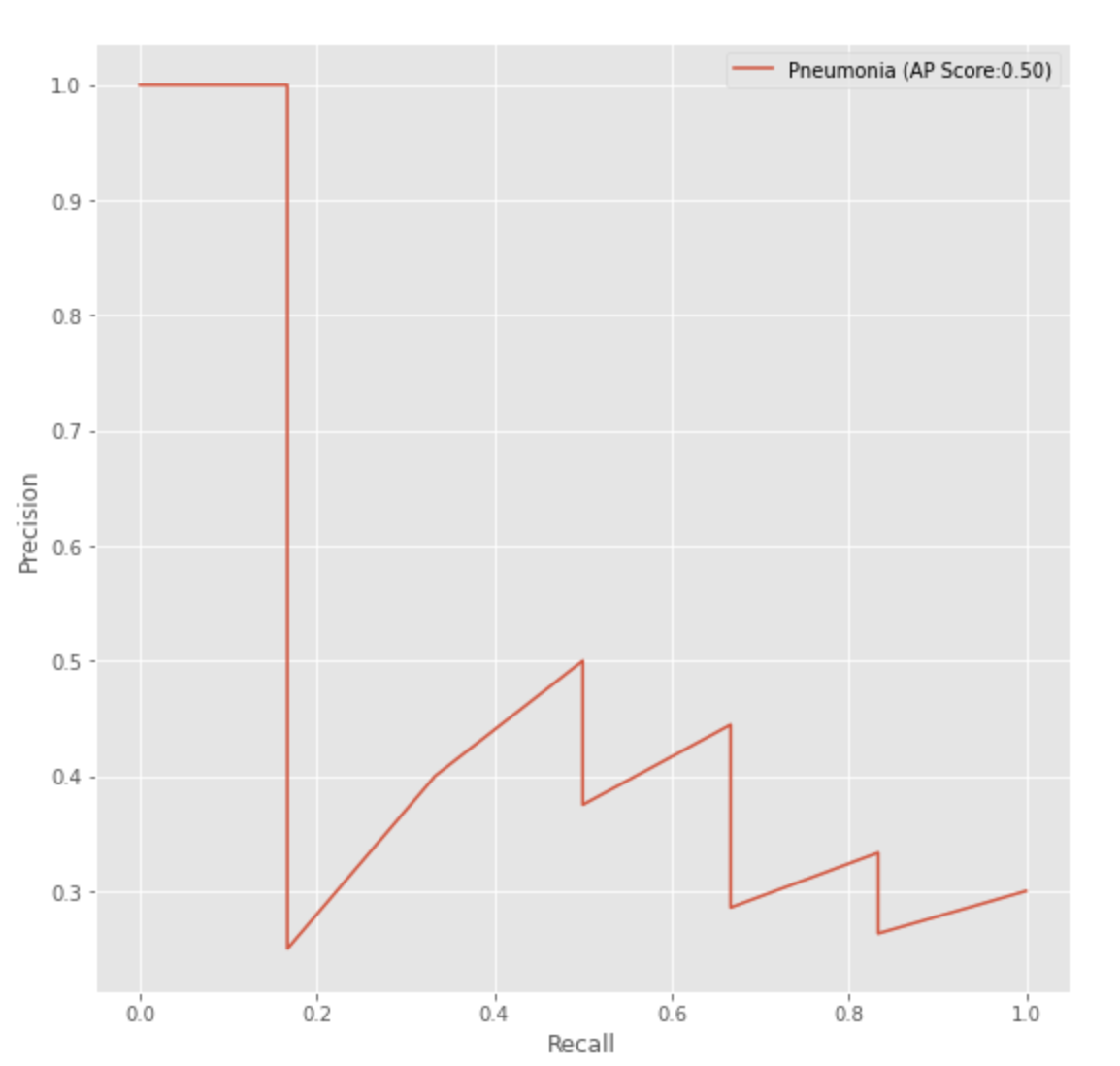
Optimizer learning rate: 1e-4

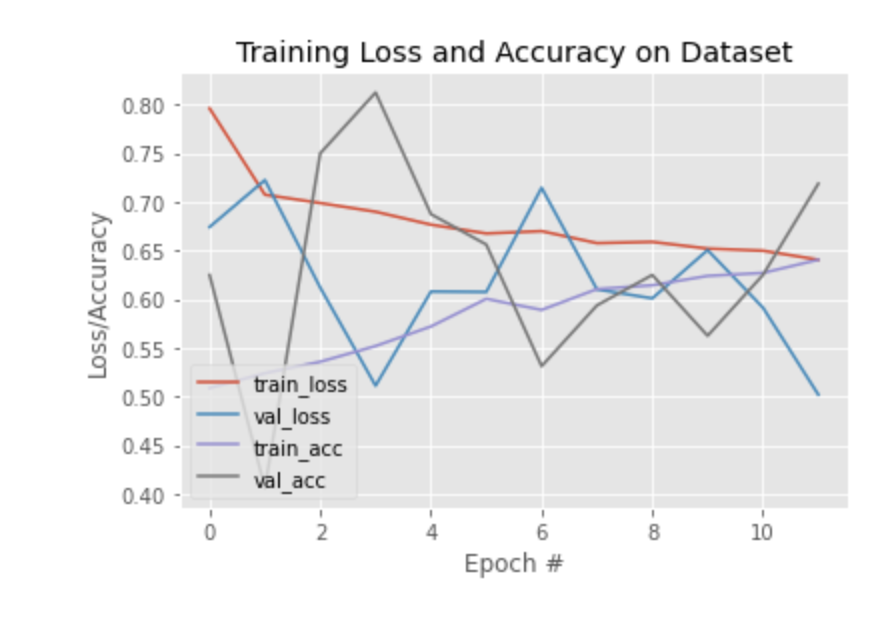
Layers of pre-existing architecture that were frozen: 17

Layers of pre-existing architecture that were fine-tuned: 1

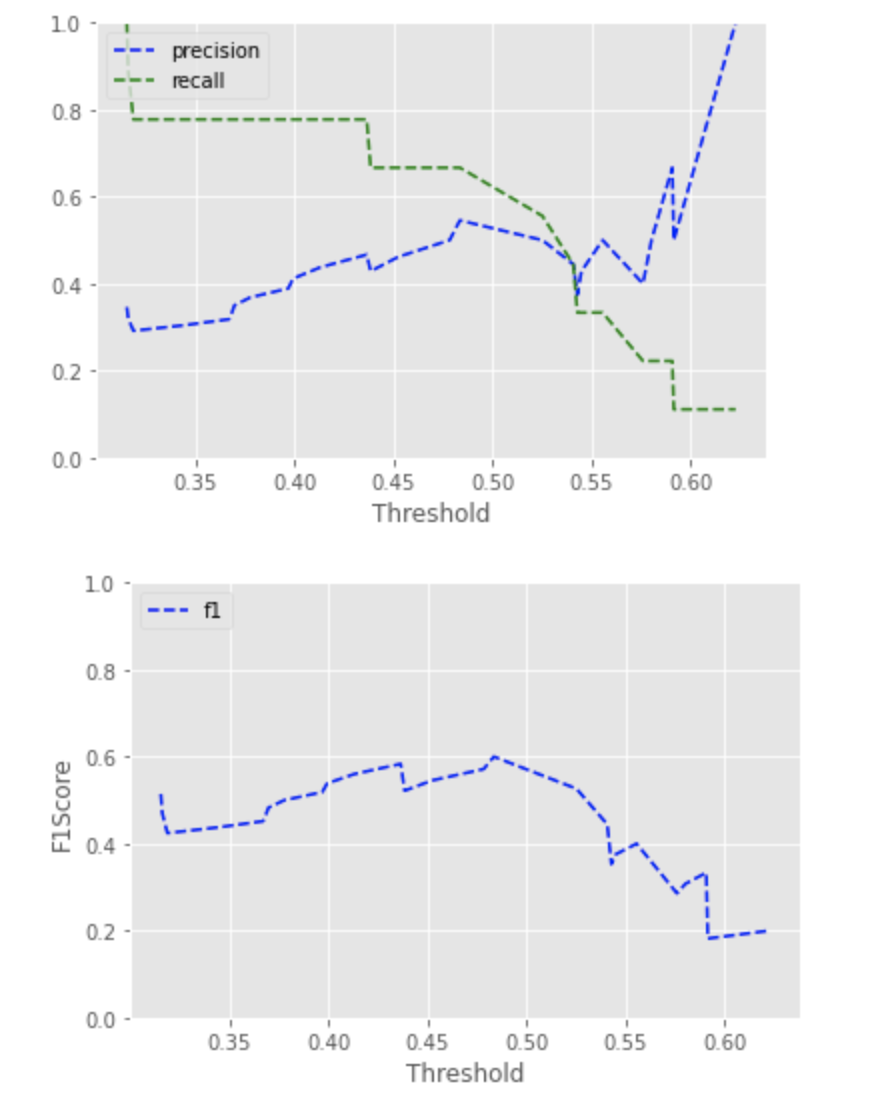
Layers added to pre-existing architecture: 3 dense fully connected layers, with 3 dropout layers







**Final Threshold and Explanation:**

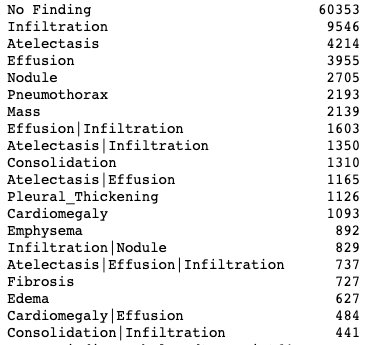
****

Threshold: 0.43. Based on the above plots i am going to take a threshold of 0.43. I want to optimize recall because that is important for screening tests. We want a low number of false negatives.

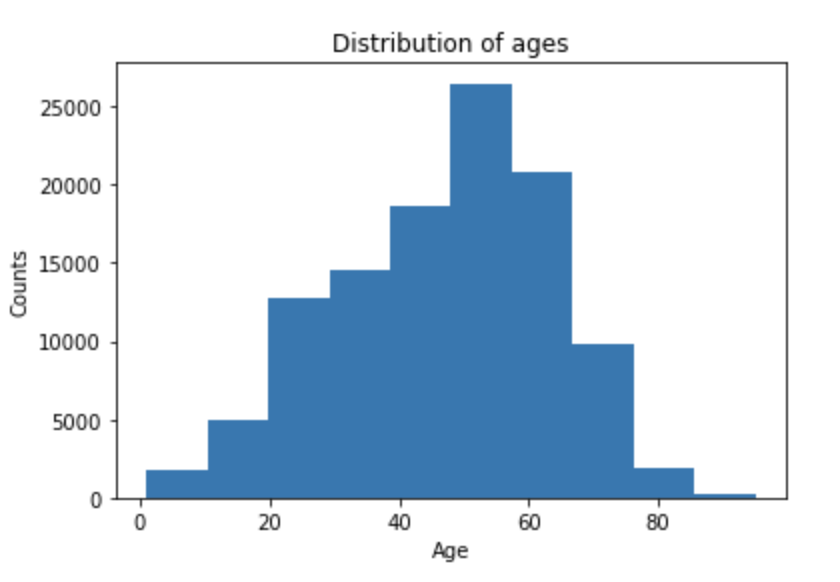
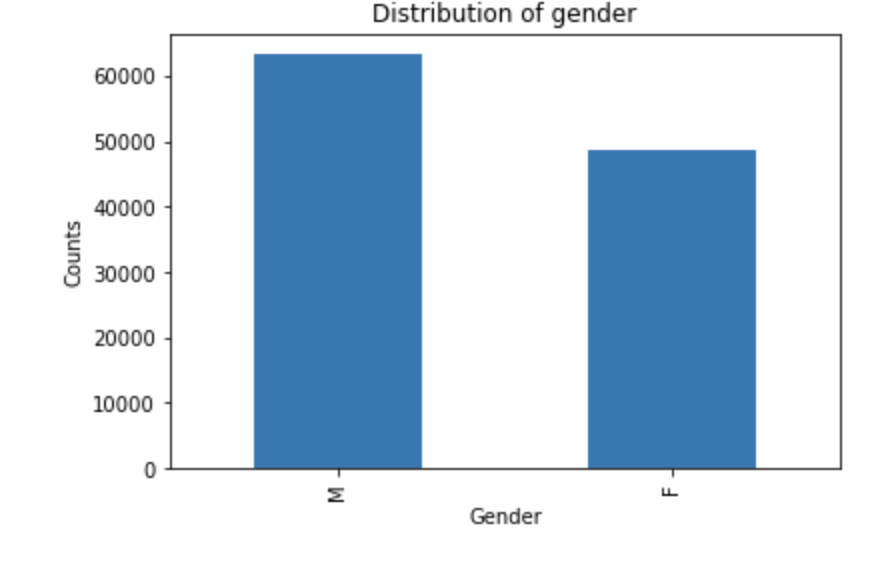
4. **Databases**

(For the below, include visualizations as they are useful and relevant)

Images were taken from the NIH Chest XRay Dataset. There are 15 unique types of labels in the dataset. The most common label is No Finding. Followed by Infiltration and Effusion





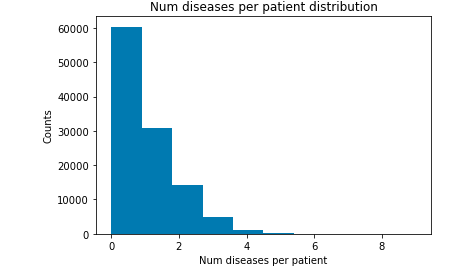


**Description of Training Dataset:**

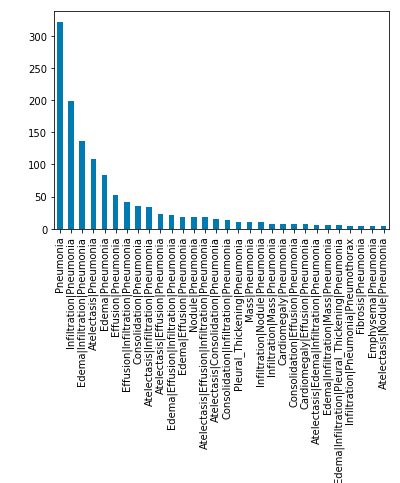
2290 Xray images with 50:50 class balance of pneumonia: no pneumonia. There are 14 other diseases that may be present in the xrays.

Xrays were taken of the chest in the PA or AP position. There are male and female patients with ages ranging 1-95 in the dataset.

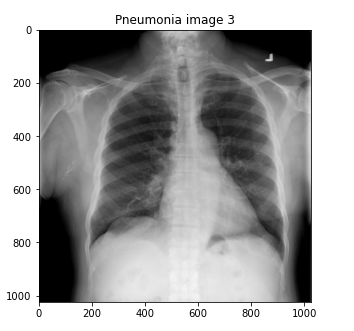
Patients most frequently have one disease but can have up to 8 diseases:



Disease combinations with pneumonia:



Pneumonia occurs most frequently alone. The most frequent comorbidities for Pneumonia are Infiltration, Edema and Atelectasis.

Example of a pneumonia xray image:  


**Description of Validation Dataset:**

1716 images with 1:5 class balance of pneumonia:no pneumonia

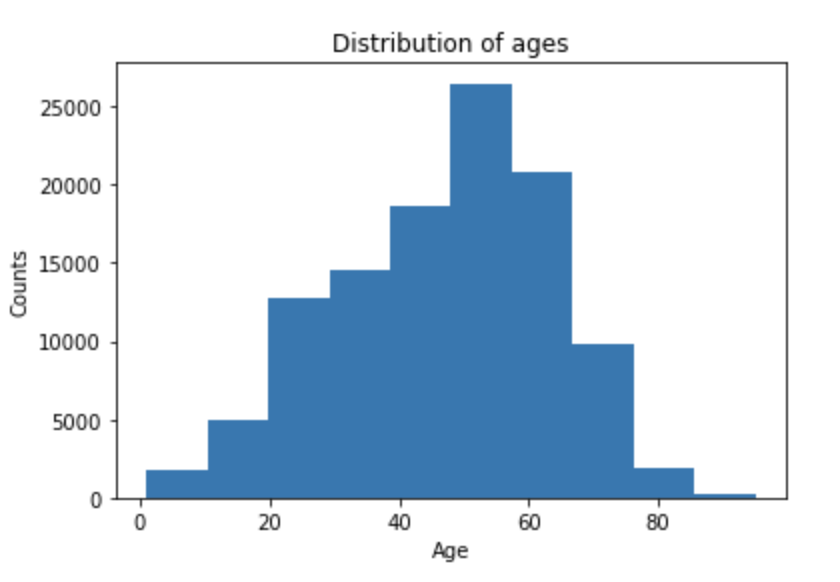
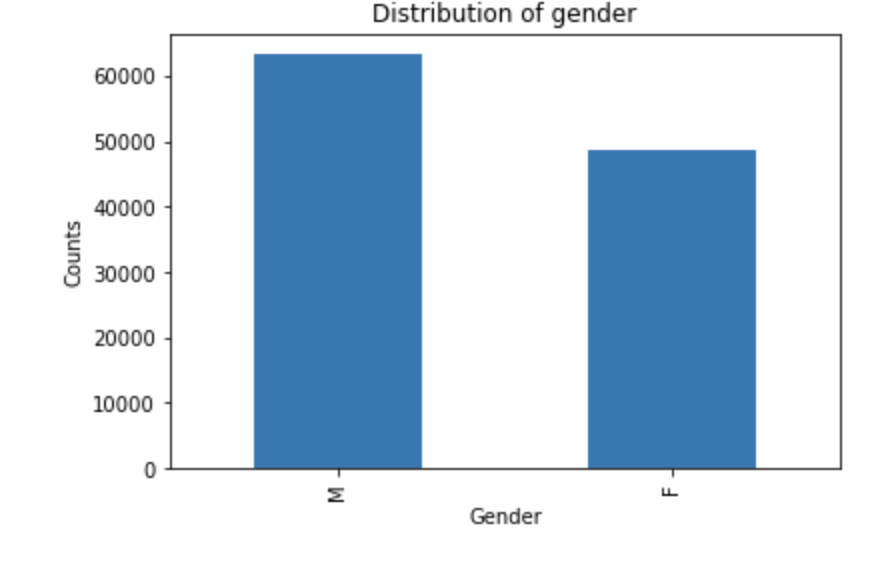
### 5. Ground Truth

Our dataset is extracted from the clinical PACS database at National Institutes of Health Clinical Center and consists of ~60% of all frontal chest x-rays in the hospital. text-mined fourteen disease image labels (where each image can have multilabels), mined from the associated radiological reports using natural language processing. The text-mined disease labels are expected to have accuracy >90%.

### 6. FDA Validation Plan

**Patient Population Description for FDA Validation Dataset:**

Male and female patients ages spanning 1 to 95 years.



Xray images of the chest are required, taken in the position PA or AP.

**Ground Truth Acquisition Methodology:**

Radiologists labels. Detecting pneumonia is hard even for trained expert radiologists, so I will use the silver standard of using several radiologists. The silver standard involves hiring *several* radiologists to each make their own diagnosis of an image. The final diagnosis is then determined by a *voting* system across all of the radiologists’ labels for each image.

**Algorithm Performance Standard:**

A highly **sensitive** test means that there are few false negative results, and thus fewer cases of disease are missed so this is the metric that I will use to measure performance. minimum acceptable sensitivity score should be 0.6 because I found in the following paper that was that was the acceptable score that they achieved [1].

# Bibliography

|  |  |
| --- | --- |
| [1] | F. T. , G. Z. ,C. C. , P. H. , S. Y. S. S. ,. K. L. ,. E. T. H. W.H. Hsu, "Development of a Deep Learning Model for Chest X-Ray Screening," *MEDICAL PHYSICS INTERNATIONAL Journal,* vol. 7, no. 3, p. 314, 2019. |