FDA Submission

Your Name: Louise O'Connor

Name of your Device: PneumonAI

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 resizes the image. Normalizing the image by subtracting the mean and dividing by the standard deviation of the image pixel intensity values.

For the model training the keras ImageDataGenerator class was used to normalize pixel values from the range of 0-255 to the range 0-1. This was achieved by setting the rescale argument to a ratio 1. / 255.0 by which each pixel can be multiplied to achieve the desired range. In this function, a random rotation range of 20 degrees was set, a height and width shift range of 0.1, and Horizontal flip. Vertical flip was set to false because you would not see an upside down Xray.

CNN Architecture:

```
Model: "model_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_1 (InputLayer)         (None, 224, 224, 3)       0
_________________________________________________________________
block1_conv1 (Conv2D)        (None, 224, 224, 64)      1792
_________________________________________________________________
block1_conv2 (Conv2D)        (None, 224, 224, 64)      36928
_________________________________________________________________
block1_pool (MaxPooling2D)   (None, 112, 112, 64)      0
_________________________________________________________________
block2_conv1 (Conv2D)        (None, 112, 112, 128)     73856
_________________________________________________________________
block2_conv2 (Conv2D)        (None, 112, 112, 128)     147584
_________________________________________________________________
block2_pool (MaxPooling2D)   (None, 56, 56, 128)       0
_________________________________________________________________
block3_conv1 (Conv2D)        (None, 56, 56, 256)       295168
_________________________________________________________________
block3_conv2 (Conv2D)        (None, 56, 56, 256)       590080
_________________________________________________________________
block3_conv3 (Conv2D)        (None, 56, 56, 256)       590080
_________________________________________________________________
block3_pool (MaxPooling2D)   (None, 28, 28, 256)       0
_________________________________________________________________
block4_conv1 (Conv2D)        (None, 28, 28, 512)       1180160
_________________________________________________________________
block4_conv2 (Conv2D)        (None, 28, 28, 512)       2359808
_________________________________________________________________
block4_conv3 (Conv2D)        (None, 28, 28, 512)       2359808
_________________________________________________________________
block4_pool (MaxPooling2D)   (None, 14, 14, 512)       0
_________________________________________________________________
block5_conv1 (Conv2D)        (None, 14, 14, 512)       2359808
_________________________________________________________________
block5_conv2 (Conv2D)        (None, 14, 14, 512)       2359808
_________________________________________________________________
block5_conv3 (Conv2D)        (None, 14, 14, 512)       2359808
_________________________________________________________________
block5_pool (MaxPooling2D)   (None, 7, 7, 512)         0
=================================================================
Total params: 14,714,688
Trainable params: 14,714,688
Non-trainable params: 0
```

```
Model: "sequential_2"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
model_1 (Model)              (None, 7, 7, 512)         14714688
_________________________________________________________________
flatten_2 (Flatten)          (None, 25088)             0
_________________________________________________________________
dense_5 (Dense)              (None, 1024)              25691136
_________________________________________________________________
dropout_4 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_6 (Dense)              (None, 512)               524800
_________________________________________________________________
dropout_5 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_7 (Dense)              (None, 256)               131328
_________________________________________________________________
dropout_6 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_8 (Dense)              (None, 1)                 257
=================================================================
Total params: 41,062,209
Trainable params: 28,707,329
Non-trainable params: 12,354,880
_________________________________________________________________
```

3. Algorithm Training

Parameters:

Types of augmentation used during training:

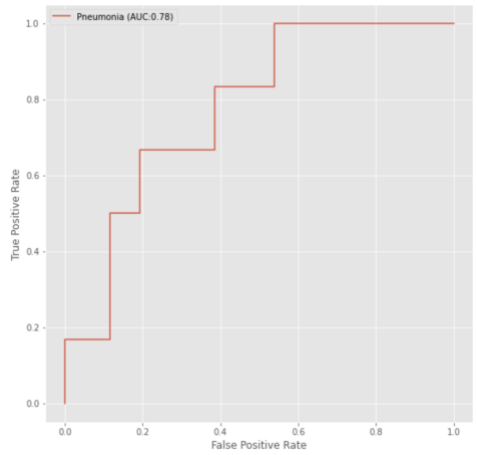
Batch size: 32

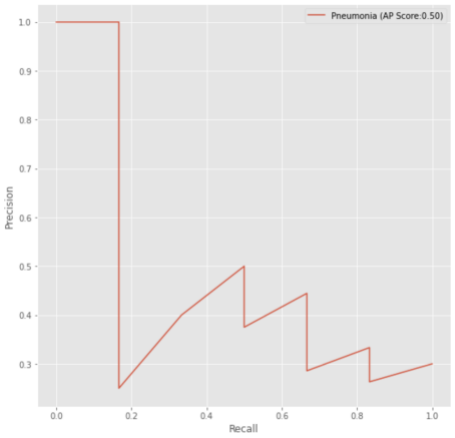
Optimizer learning rate: 1e-4

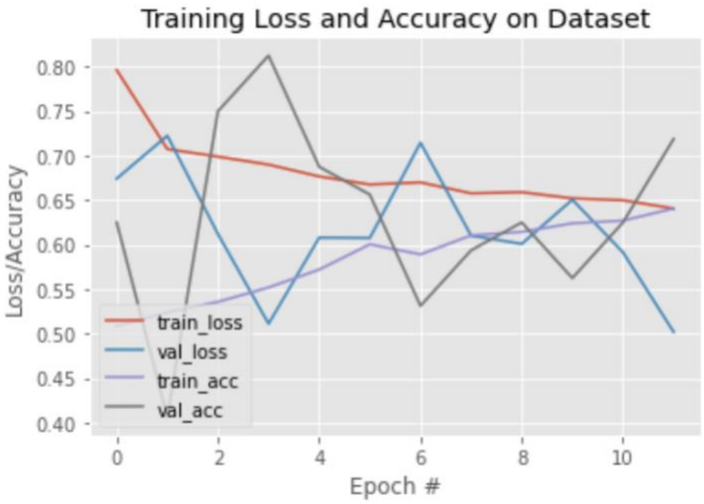
Layers of pre-existing architecture that were frozen: 17

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

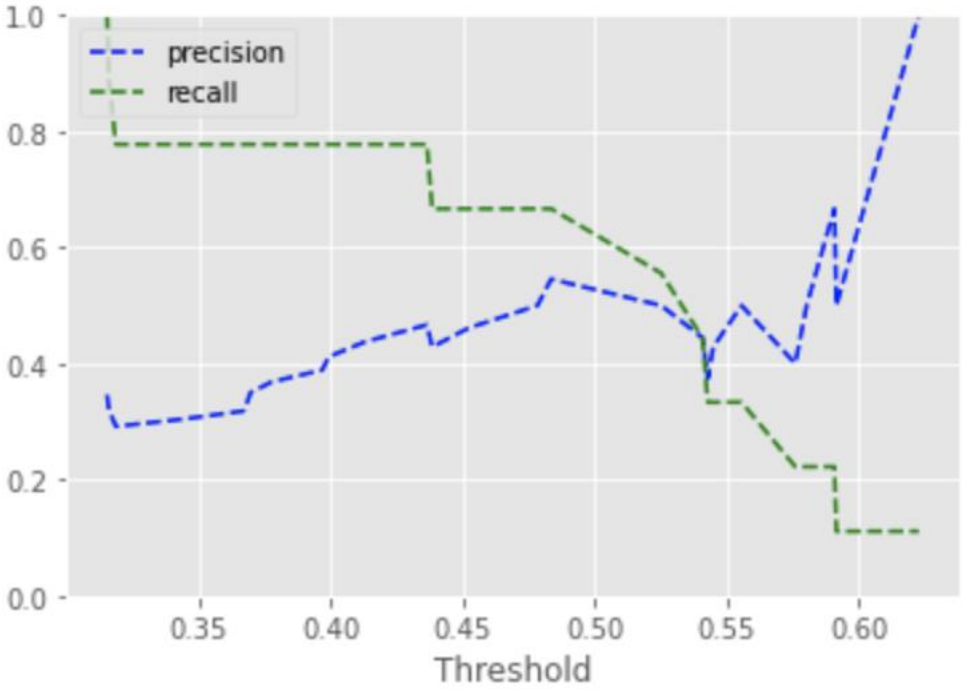
Layers added to pre-existing architecture: 3 dense fully connected layers, with 3 dropout layers

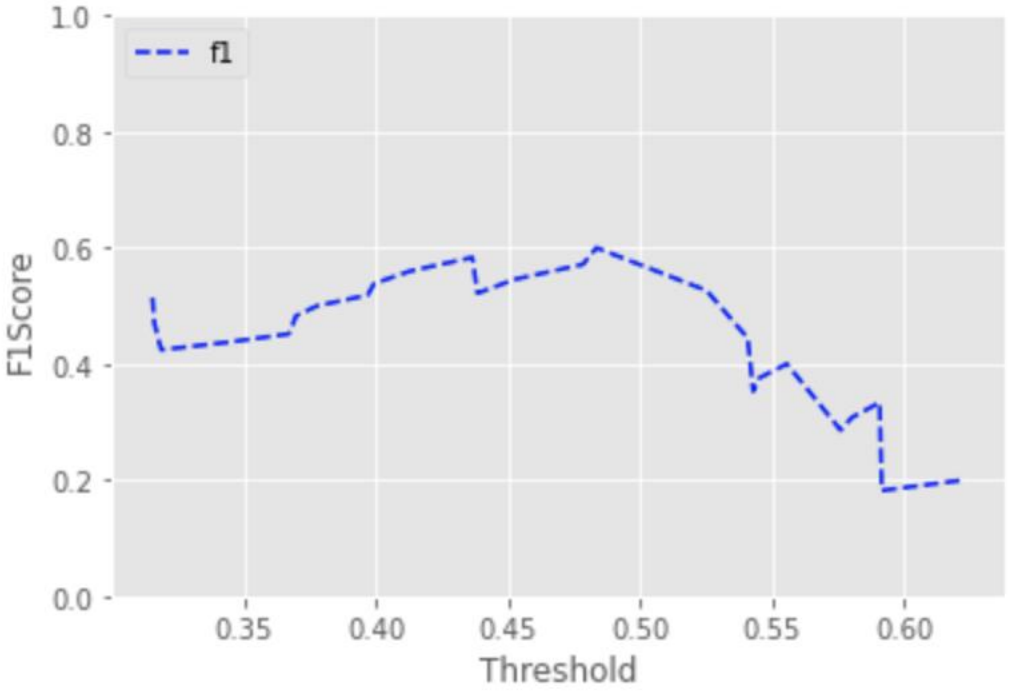

Pneumonia (AUC:0.78)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0


Pneumonia (AP Score:0.50)
Precision
Recall
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


Training Loss and Accuracy on Dataset
Loss/Accuracy
Epoch #
train_loss
val_loss
train_acc
val_acc
0.40
0.45
0.50
0.55
0.60
0.65
0.70
0.75
0.80
0
2
4
6
8
10

Final Threshold and Explanation:





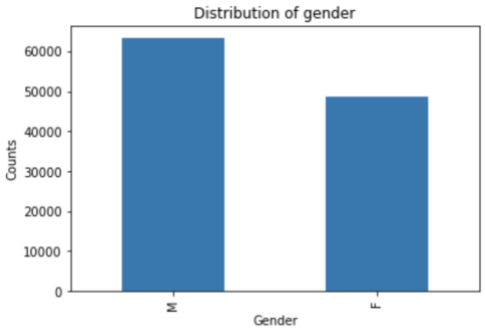
Threshold: 0.43. Based on the above plots i am going to take a threshold of 0.43. I want to optimize recall because that is important for screening tests. We want a low number of false negatives.

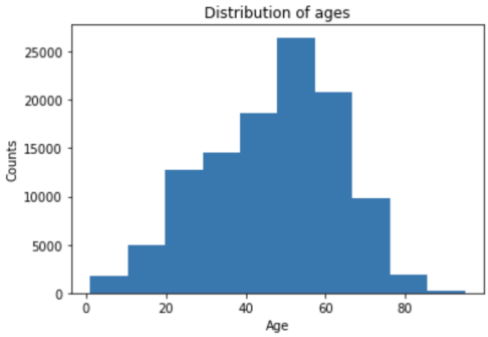
4. **Databases**

(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

```
No Finding                          60353
Infiltration                         9546
Atelectasis                          4214
Effusion                             3955
Nodule                               2705
Pneumothorax                         2193
Mass                                 2139
Effusion|Infiltration                1603
Atelectasis|Infiltration             1350
Consolidation                        1310
Atelectasis|Effusion                 1165
Pleural_Thickening                   1126
Cardiomegaly                         1093
Emphysema                             892
Infiltration|Nodule                   829
Atelectasis|Effusion|Infiltration     737
Fibrosis                              727
Edema                                 627
Cardiomegaly|Effusion                 484
Consolidation|Infiltration            441
Number of cases: 112104
Number of pneumonia cases: 1430
Number of pneumonia cases: 110674
```

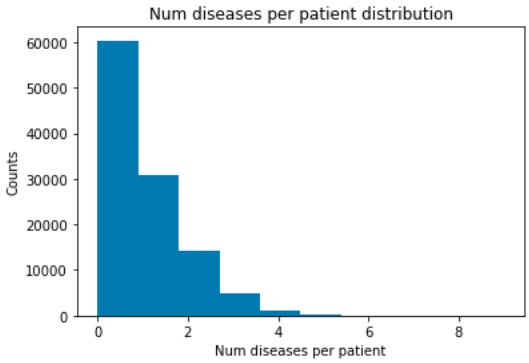




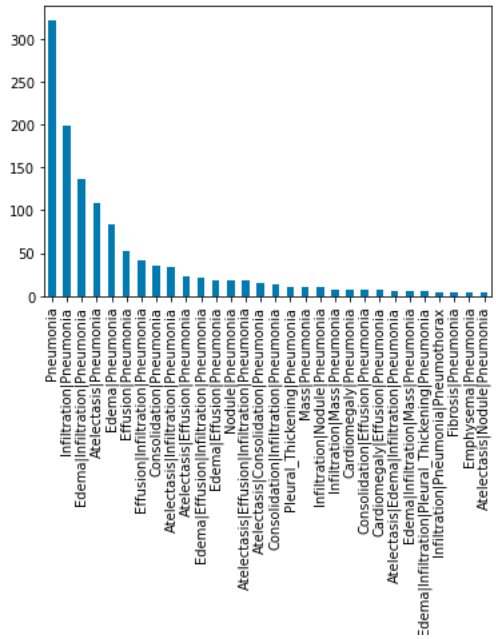
Description of Training Dataset:

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:

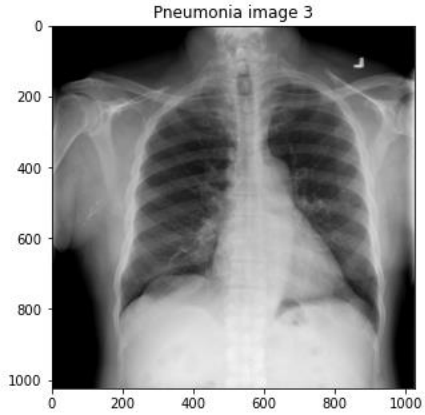


Disease combinations with pneumonia:



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 i.e. 286 images with pneumonia, 1430 images without pneumonia, Images are chest xrays in position AP or PA, with male and female patients, ages ranging 1-95 years.

Patients most commonly have 1 disease but may have up to 8 diseases or 'No Finding'

The following list is of the 15 disease labels: ['Atelectasis', 'Cardiomegaly', 'Consolidation', 'Edema', 'Effusion', 'Emphysema', 'Fibrosis', 'Hernia', 'Infiltration', 'Mass', 'No Finding', 'Nodule', 'Pleural_Thickening', 'Pneumonia', 'Pneumothorax']

The top 20 value counts of the finding labels:

No Finding	792
Infiltration	129
Pneumonia	66
Effusion	47
Atelectasis	46
Infiltration|Pneumonia	39
Pneumothorax	36
Edema|Infiltration|Pneumonia	29
Nodule	27
Mass	26
Effusion|Infiltration	25
Atelectasis|Pneumonia	25
Consolidation	18
Cardiomegaly	18
Effusion|Pneumonia	14
Edema|Pneumonia	13
Pleural_Thickening	13
Emphysema	12
Infiltration|Nodule	11
Cardiomegaly|Effusion	11

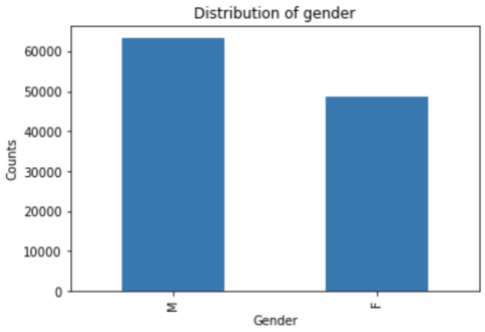
5. Ground Truth

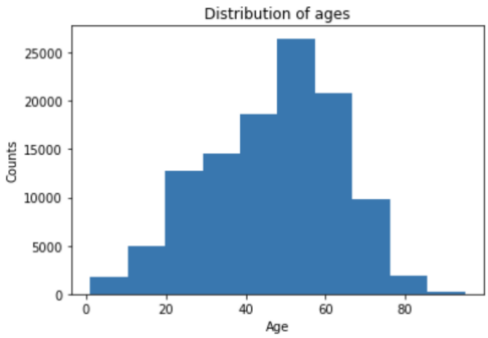
Our dataset is extracted from the clinical PACS database at National Institutes of Health Clinical Center and consists of ~60% of all frontal chest x-rays in the hospital. text-mined fourteen disease image labels (where each image can have multilabels), mined from the associated radiological reports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