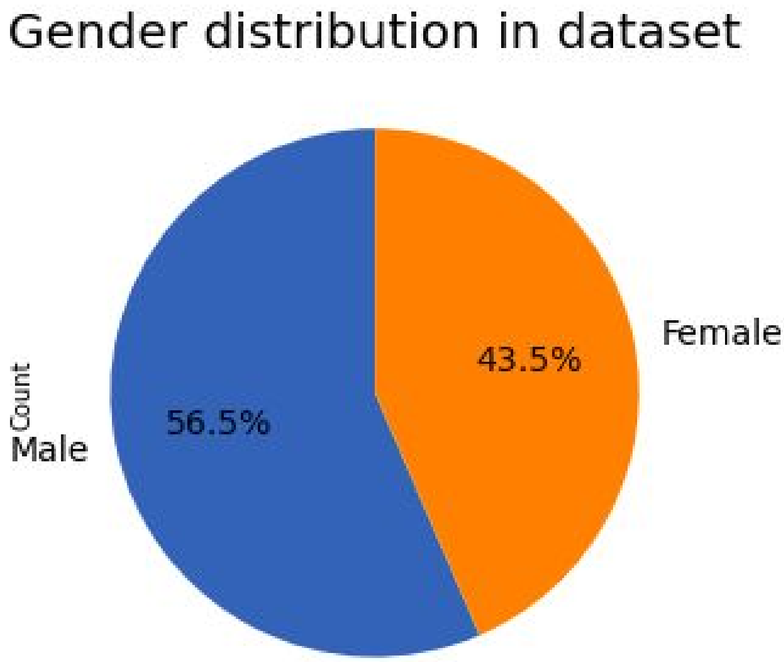
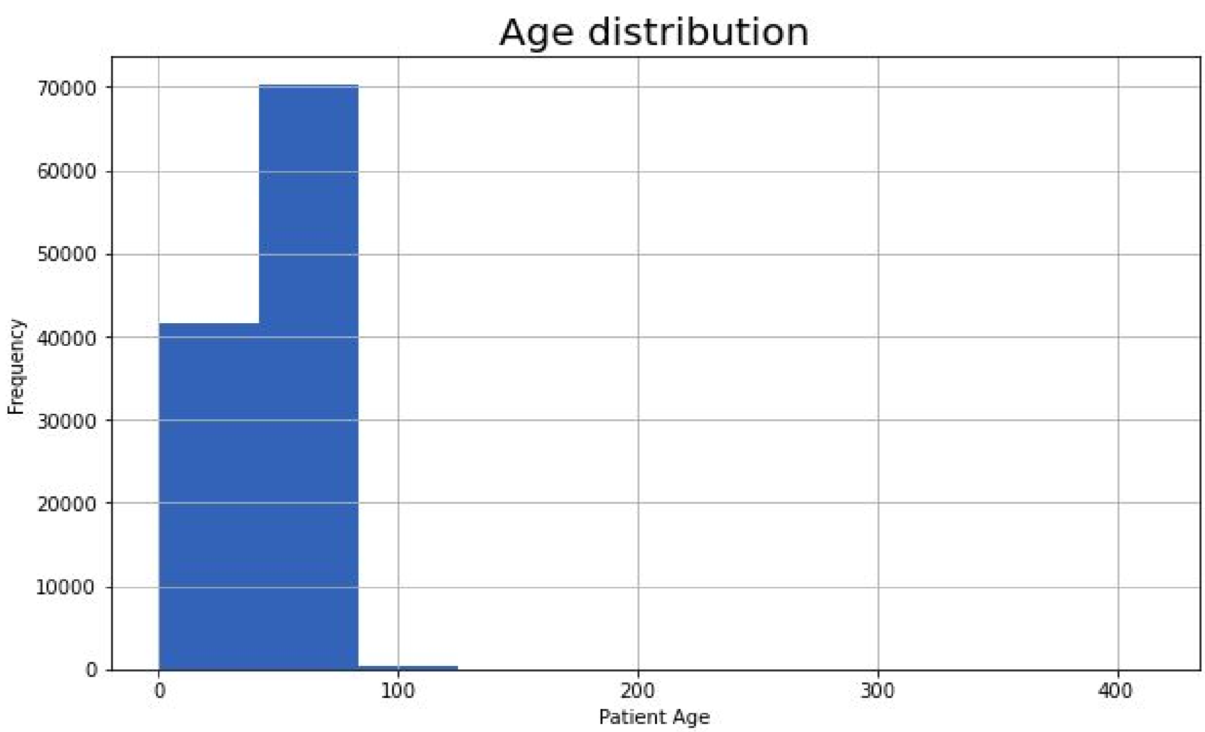
FDA Submission

Name: Rahul

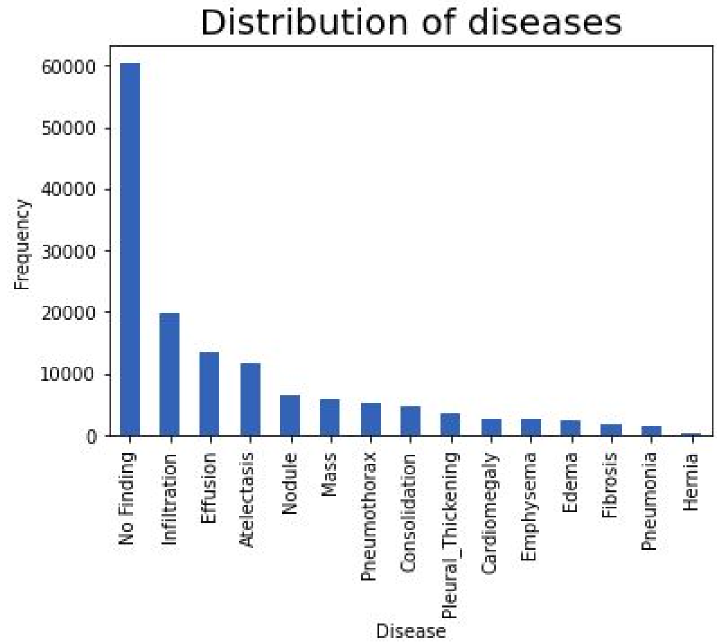
Device Name: Pneumonia Detector

General Information





```
Pneumonia case: 1431
```



Intended Use Statement:
Looking at the above, we see that the algorithm was trained on 56.5% male and 43.5% female patients who spanned from ages 1 to 414. All patients were scanned for the chest X-ray and were labeled with 14 diseases and No finding.

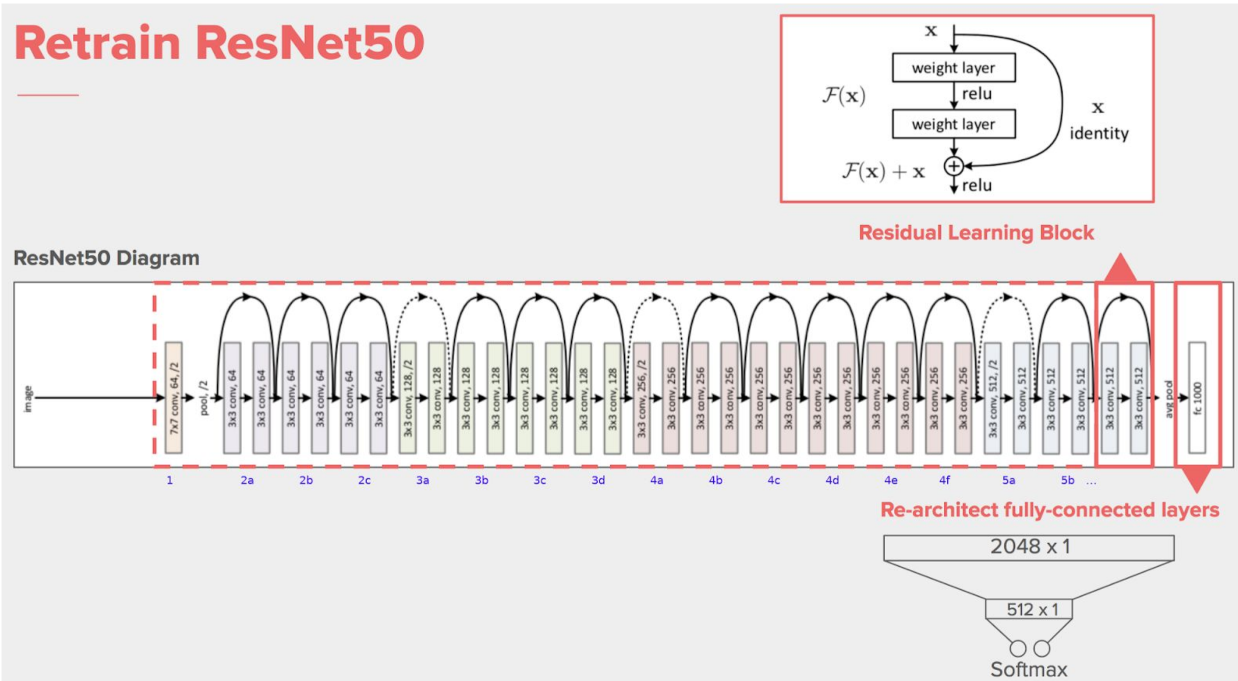
From this information, the appropriate intended use statement would be:
This algorithm is intended for use on Pneumonia patients from the ages of 1-100 (as there are outliers in the dataset) who have been administered a screening chest X-ray and have never before demonstrated an abnormal Chest X-ray study.

Indication of Use:
The algorithm can be used for the screening of the chest X-ray which can be helpful in the early detection of Pneumonia.

Device Limitations:
The results of the algorithm indicate that the presence of Atelectasis, Cardiomegaly, Consolidation, Edema, Effusion, Emphysema, Fibrosis, Hernia, Infiltration, Mass, Nodule, Pleural Thickening, or Pneumothorax in a chest x-ray may lead to false-positive pneumonia classifications as specificity is 1, however, the presence of Pneumonia can be accurately detected from this algorithm.

Clinical Impact of Performance:
The algorithm has a very good result in detecting Pneumonia which can be very beneficial in automating the process of detecting Pneumonia cases directly from the chest X-ray which not saves a lot of time in diagnosing the patient and also preparing the patient mentally as well.

Algorithm Design and Function

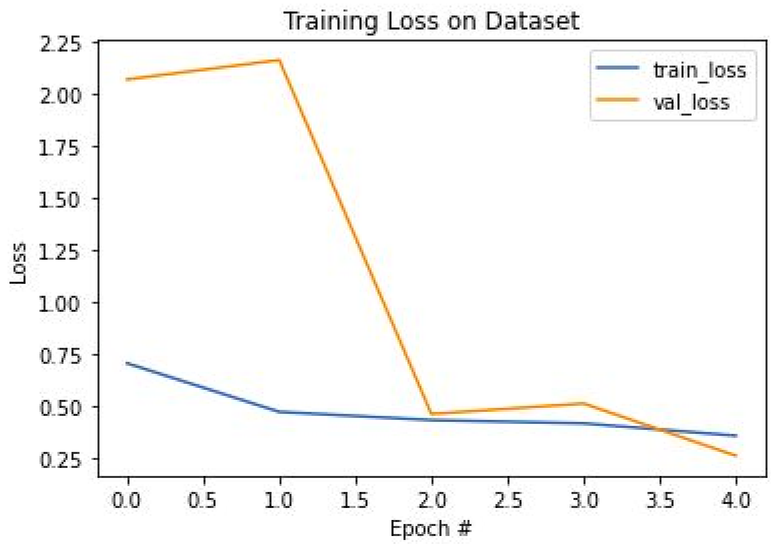


The algorithm is designed based on the pretrained resnet50 model. After the last layer of the resnet model. Following layers are added:

```
my_model = Sequential([resnet_model,
                       BatchNormalization(),
                       Conv2D(1024, 1, activation='relu'),
                       Dropout(0.5),
                       BatchNormalization(),
                       Conv2D(256, 1, activation='relu'),
                       Dropout(0.5),
                       AveragePooling2D((7,7)),
                       BatchNormalization(),
                       Conv2D(1, 1, activation='sigmoid'),
                       Reshape((-1,))
                      ])
```

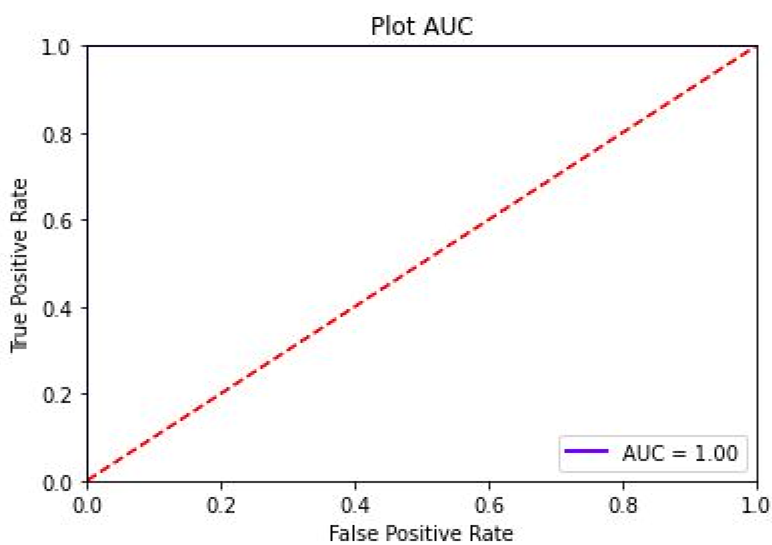
For training data augmentation, the image is rescaled to 1/255, horizontal flip, height shift range of 0.1, width shift range of 0.1, rotation range of 20, shear range of 0.1, the zoom range of 0.1, so that model will consider those image as well which are slightly rotated or zoomed. For the training image generator, a batch size of 64 images is used.

The algorithm is trained based on RMSprop optimizer and a learning rate of 10^-4. Then the model remaining is fine-tuned for this case for about more than 50 epochs until the model loss is stabilized.



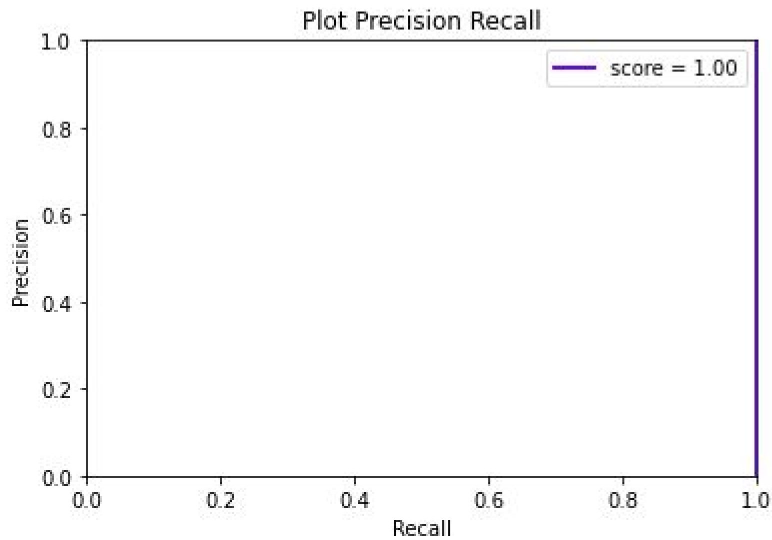
**

The current model was trained up to 98.21% training accuracy and training loss of 4.89% which has a final validation accuracy of 98.438%.



**

The algorithm has an area under the curve for True positive rate and false positive rate of 1.



**

And the precision-recall curve has an AP score of 1.

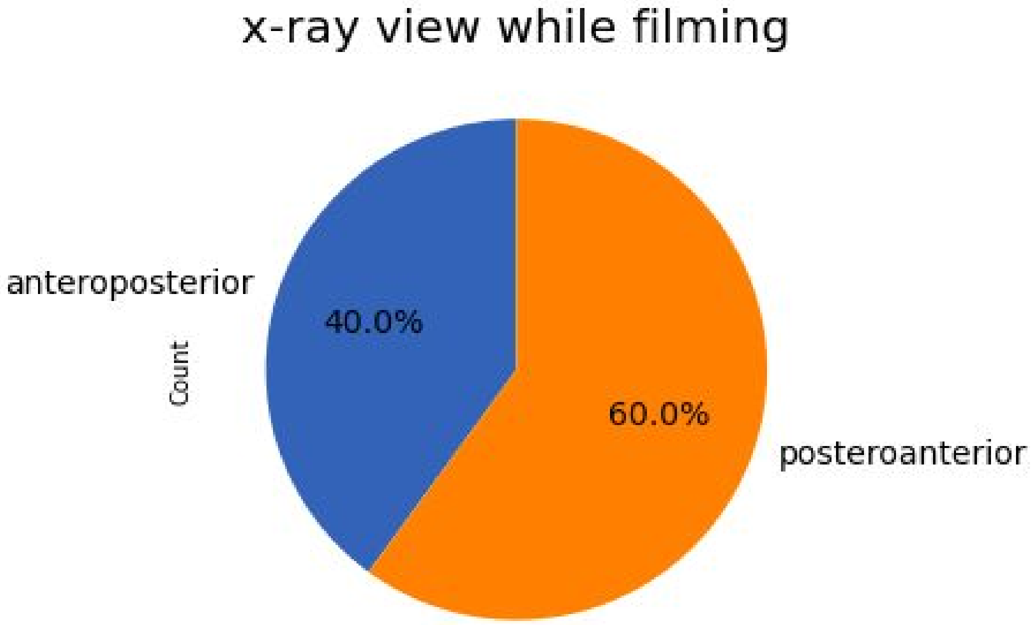
```
        Threshold of 1.00 gives best accuracy at 0.9688
Out[18]: 0.888888888888889
```

From the above, the calculated threshold is `1.00` with the F1 score of `0.889` which has the best accuracy of 96.88%`.`

Databases

The dataset was curated by the NIH. There are 112,120 X-ray images with disease labels from 30,805 unique patients in this dataset. The disease labels were created using Natural Language Processing (NLP) to mine the associated radiological reports. The labels include 14 common thoracic pathologies:

- Atelectasis
- Consolidation
- Infiltration
- Pneumothorax
- Edema
- Emphysema
- Fibrosis
- Effusion
- Pneumonia
- Pleural thickening
- Cardiomegaly
- Nodule
- Mass
- Hernia



The dataset consists of 56.5% male participants and 43.5% female participants (can be depicted from the pie chart on the first page). The participants in the dataset range from 1-100 with an outlier at 414. The positioning of the participant during the chest X-ray is AP in 40% cases and PA during 60% cases. There are 1413 pneumonia cases in the given dataset which can be in the presence of different 15 images. The overall distribution of the diseases can be found from the first image of the 2nd page.

For training the model, I have split the dataset into 80% training data and a 20% dataset. Then the training data is balanced for 50% pneumonia cases and 50% non-pneumonia cases.
The training dataset is augmented as described in the previous section then the training data frame is generated with the target size of 224 x 224 and a batch size of 64.
Similarly, the validation dataset is generated as the training dataset but with image augmentation ie just rescaling to 1/255.

Ground Truth

The biggest limitation of this dataset is that image labels were NLP-extracted so there could be some erroneous labels but the NLP labeling accuracy is estimated to be >90%.

The original radiology reports are not publicly available but details about the labeling process can be found here.

FDA Validation Plan

Patient Population Description for FDA Validation Dataset:
For validation of the algorithm, the collected dataset should be made up of chest X-rays between the ages 1-100 for both male and female. However, it should be made sure that the validation set did not contain a patient with prior history of Pneumonia and the patient with other diseases like Atelectasis, Cardiomegaly, Consolidation, Edema, Effusion, Emphysema, Fibrosis, Hernia, Infiltration, Mass,

Nodule, Pleural Thickening, or Pneumothorax should be checked again with Radiologists for False Positive.

Ground Truth Acquisition Methodology:

For the acquisition of the chest X-ray, the patient can be in any position ie anteroposterior or posteroanterior. The patient should also be checked other diseases mentioned above, as it may result in False positive.

Algorithm Performance Standard:

The silver standard approach of using several radiologists would be more optimal for this Algorithm.

**Note: All the training works can be found here.