# FDA Submission

## Claire Vania

**Name of the Device:** PneumonoTec

**Algorithm Description:**

### 1. General Information

**Intended Use Statement:** Assisting radiologist in detection of pneumonia from Xray

**Indications for Use:** Intended to use on men and women in the age of 1 to 85 with view position being AP or PA

**Device Limitations:**

- Prediction time depends upon the processing speed of CPU
- Presence of Infiltration or Atelectasis can affect the prediction

**Clinical Impact of Performance**: The model had lower precision and higher recall which means that there would be more false positives than false negatives and hence the model should only be used to assist clinicians and not to conclude the presence or absence of a disease

### 2. Algorithm Design and Function


**DICOM Checking Steps:**

- Body Part: Chest
- Position: PA or AP
- Modality: DX

**Preprocessing Steps:**

- Reshaping
- Normalizing
- Repeating the process for all three channels

**CNN Architecture:** We use a pretrained VGG16 model. The VGG16 model layers were frozen except the last convolution layer and then flattened to pass through fully connected layers

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
model_2 (Model)              (None, 7, 7, 512)         14714688
_________________________________________________________________
flatten_1 (Flatten)          (None, 25088)             0
_________________________________________________________________
dropout_1 (Dropout)          (None, 25088)             0
_________________________________________________________________
dense_1 (Dense)              (None, 1024)              25691136
_________________________________________________________________
dropout_2 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_2 (Dense)              (None, 512)               524800
_________________________________________________________________
dropout_3 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_3 (Dense)              (None, 256)               131328
_________________________________________________________________
dense_4 (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:**

- shear_range = 0.1,
- zoom_range = 0.1,
- rotation_range = 20,
- height_shift_range = 0.1,
- width_shift_range = 0.1

**Batch size:** 32

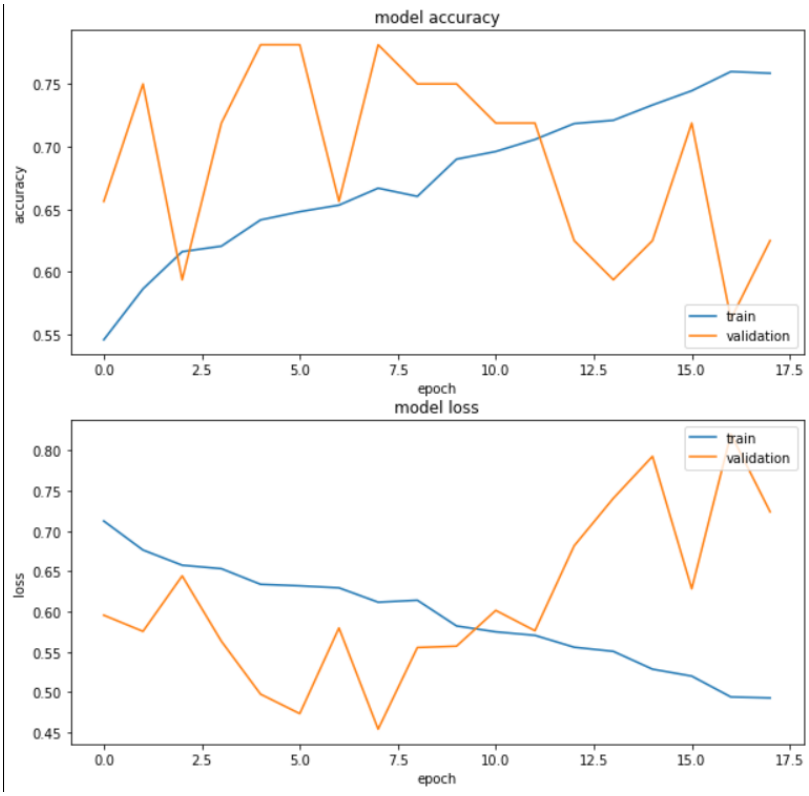
**Optimizer learning rate:** 1e-4

**Layers of pre-existing architecture that were frozen:** 17 layers of VGG Model

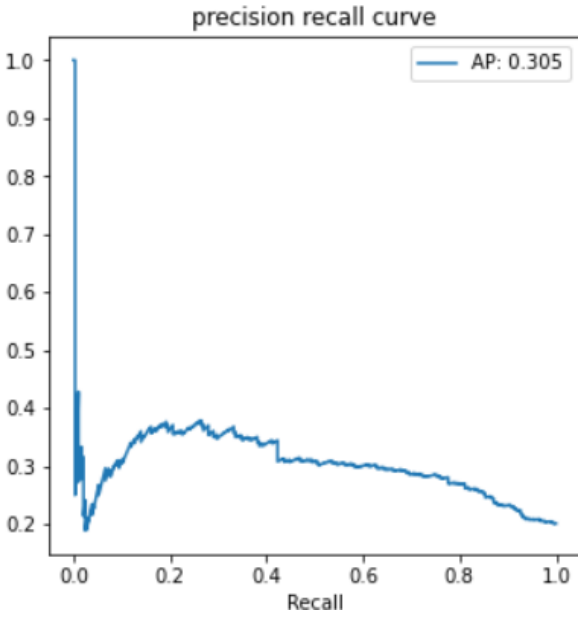
**Layers of pre-existing architecture that were fine-tuned:** Last two layers

**Layers added to pre-existing architecture:** flatten, dropout and dense layers as shown in the architecture

**Algorithm Training Performance Visualization:**



**P-R curve**

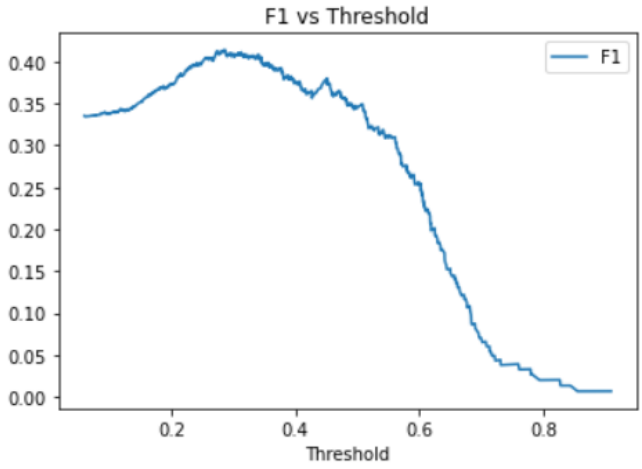



**Final Threshold:** Threshold was selected where F1 score was maximum

The F1 score from the paper CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.

| | F1 Score (95% CI) |
|---|---|
| Radiologist 1 | 0.383 (0.309, 0.453) |
| Radiologist 2 | 0.356 (0.282, 0.428) |
| Radiologist 3 | 0.365 (0.291, 0.435) |
| Radiologist 4 | 0.442 (0.390, 0.492) |
| Radiologist Avg. | 0.387 (0.330, 0.442) |
| CheXNet | 0.435 (0.387, 0.481) |

```
Maximum f1: 0.41352657004830917
Threshold: 0.28675130009651184
```



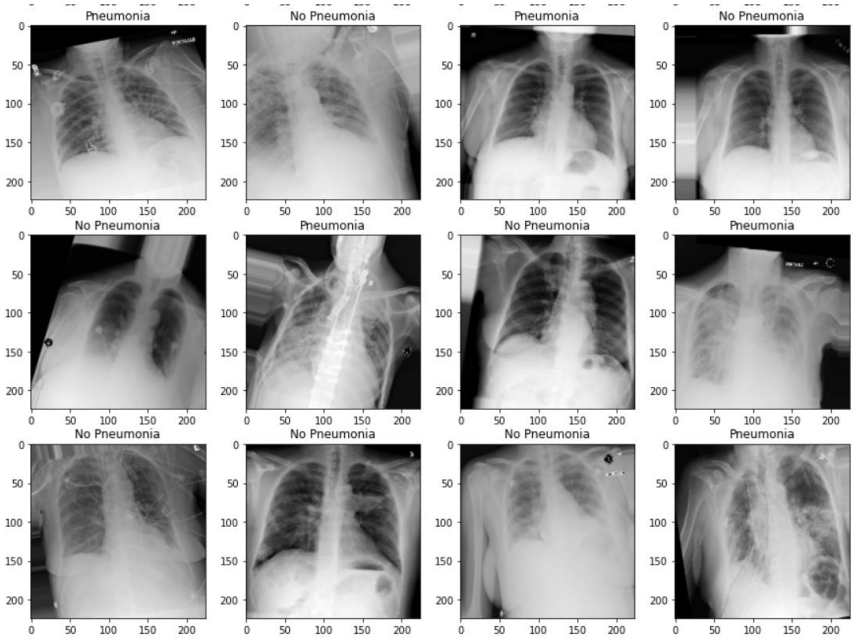

The models F1 score is better than most radiologist and comparable to that of CheXNet.

## 4. Databases

NIH Chest X-ray Dataset was used which includes 112000 chest xrays with disease labels acquired from 30000 patients.

### Description of Training Dataset:

The training set has equal population of pneumonia and non-pneumonia samples with total samples being 2290. Some training samples.



**Description of Validation Dataset:**

Validation dataset had 1:4 ratio of pneumonia positive to pneumonia negative cases to represent real world scenario. Total samples are 1430

## 5. Ground Truth

There are 112,120 X-ray images with disease labels from 30805 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.

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

The biggest limitation of the 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%

## 6. FDA Validation Plan

**Patient Population Description for FDA Validation Dataset:** Patients can be male or female in the age range of 1 – 85 with chest xray being taken in either PA or AP position. Since the dataset should high correlation between infiltration and Atelectasis with pneumonia, the validation set should not include patients with these diseases.

**Ground Truth Acquisition Methodology:** The gold standard is by obtaining a lab report for the presence or absence of pneumonia bacteria. Since this is time consuming, we can rely on silver standard which is taking weighted sum of several radiologist defined labels.

**Algorithm Performance Standard:** The algorithms F1 score should be more than that of an average radiologists(0.387) as mentioned in the ChesXNet paper.