

FDA Submission

Your Name: Shyam Panjwani

Name of your Device: Pneumonia Detector using 2D Xray Images

Algorithm Description

1. General Information

Intended Use Statement:

This device should be used to assist radiologists and other doctors in detecting Pneumonia in patients.

Indications for Use:

Since this device was built using data from both Male and Female patients with age ranging from 0 to 100. The training dataset also has both positions 'PA' and 'AP'. This device can be used for wide range of patients.

Device Limitations:

The device may not produce very accurate results in case patient has other diseases as well.

Clinical Impact of Performance:

The device can be used to prioritize treatments of patients which show Pneumonia symptoms as per the device prediction.

2. Algorithm Design and Function

DICOM Checking Steps:

Meta information like Patient Age, Gender etc. are checked in DICOM files. Pixel image data are extracted from the DICOM files

Preprocessing Steps:

- 1) Pixel array obtained from DICOM file are normalized (intensities divided by 255).
- 2) Pixel images are resized to (224,224)
- 3) Pixel images could be mean centered and normalized by std. deviation. However, this step was not implemented because mean centering and normalization produced worse validation results.

****CNN Architecture:****

Layer (type)	Output Shape	Param #
model_3 (Model)	(None, 7, 7, 512)	14714688
flatten_3 (Flatten)	(None, 25088)	0
dropout_7 (Dropout)	(None, 25088)	0
dense_9 (Dense)	(None, 1024)	25691136
dropout_8 (Dropout)	(None, 1024)	0
dense_10 (Dense)	(None, 512)	524800
dropout_9 (Dropout)	(None, 512)	0
dense_11 (Dense)	(None, 256)	131328
dense_12 (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

rotation_range=20,
width_shift_range=0.1,
height_shift_range=0.1,
shear_range=0.1,
zoom_range=0.1,
horizontal_flip=True,
vertical_flip=False,
rescale=1.0/255

* Batch size: 32

* Optimizer learning rate: $1e-4$

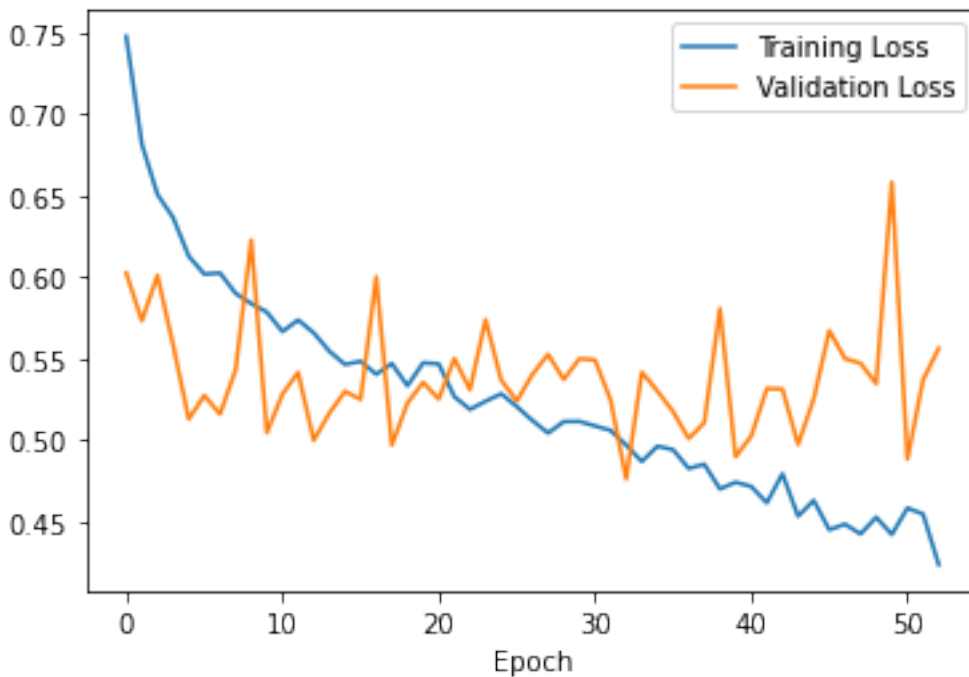
* Layers of pre-existing architecture that were frozen: 18 layers of VGG16 model were frozen

* Layers of pre-existing architecture that were fine-tuned: None

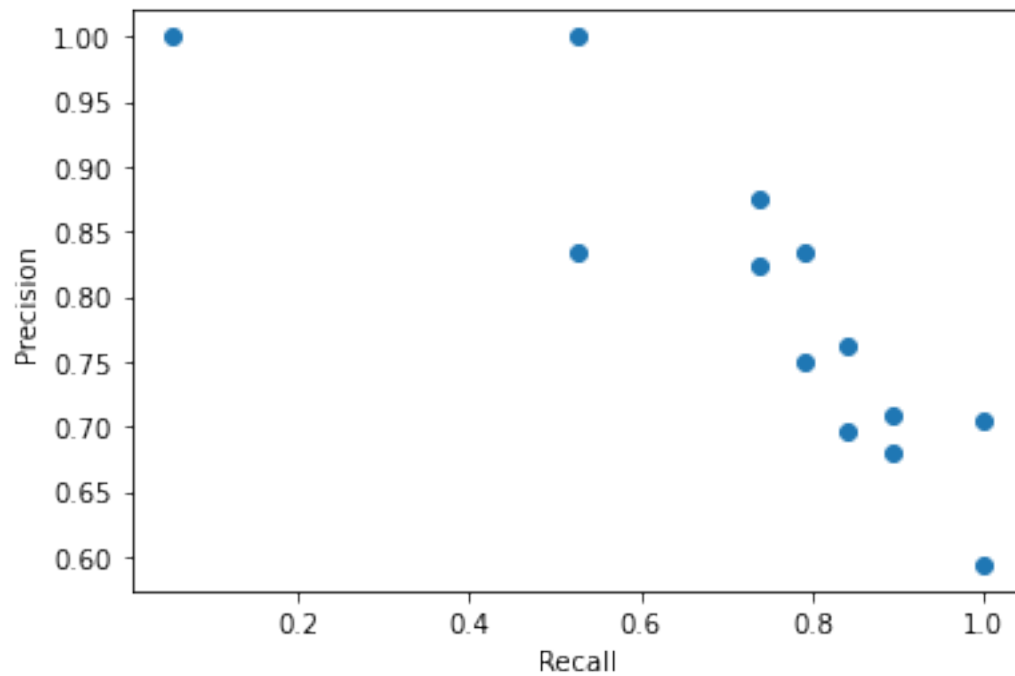
* Layers added to pre-existing architecture:

- 1) Flatten layer
- 2) Dropout (3 layers)
- 3) Dense (3 layers)
- 4) Dense (1 output layer)

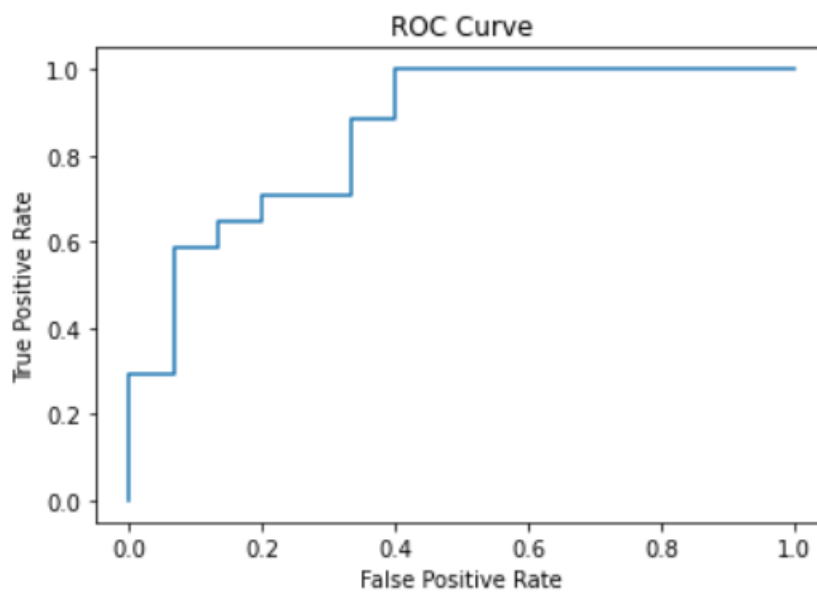
<< Insert algorithm training performance visualization >>



<< Insert P-R curve >>



Roc Curve



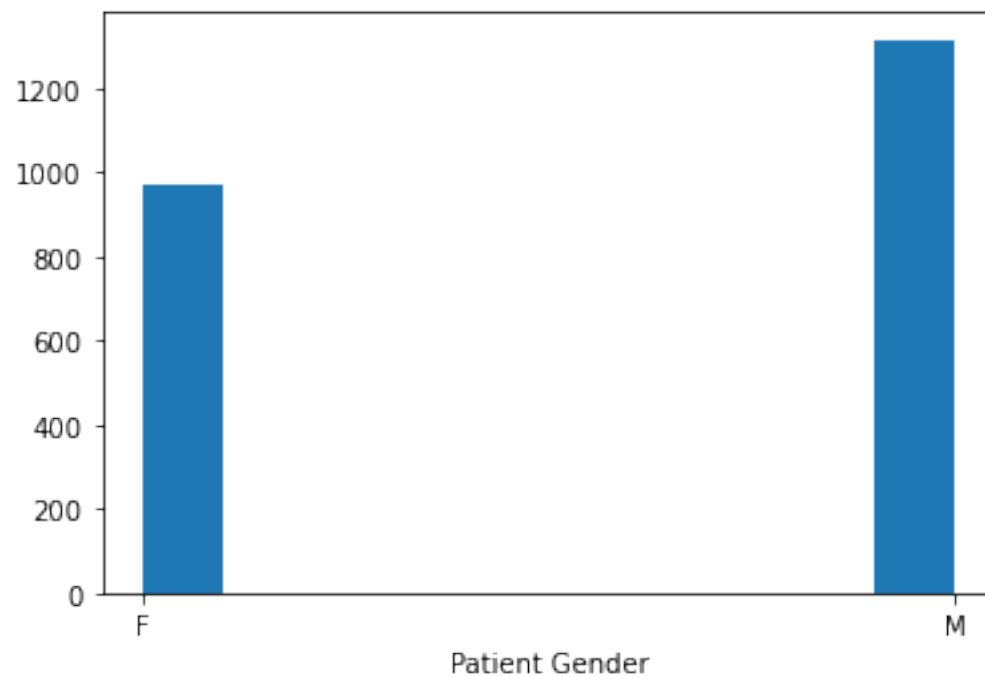
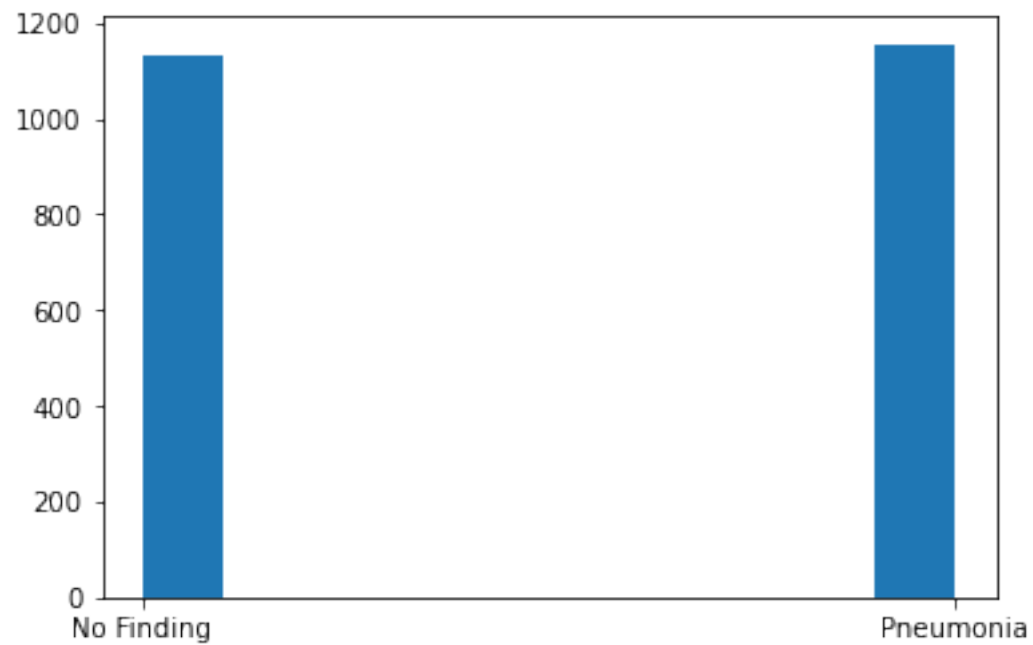
Final Threshold and Explanation:

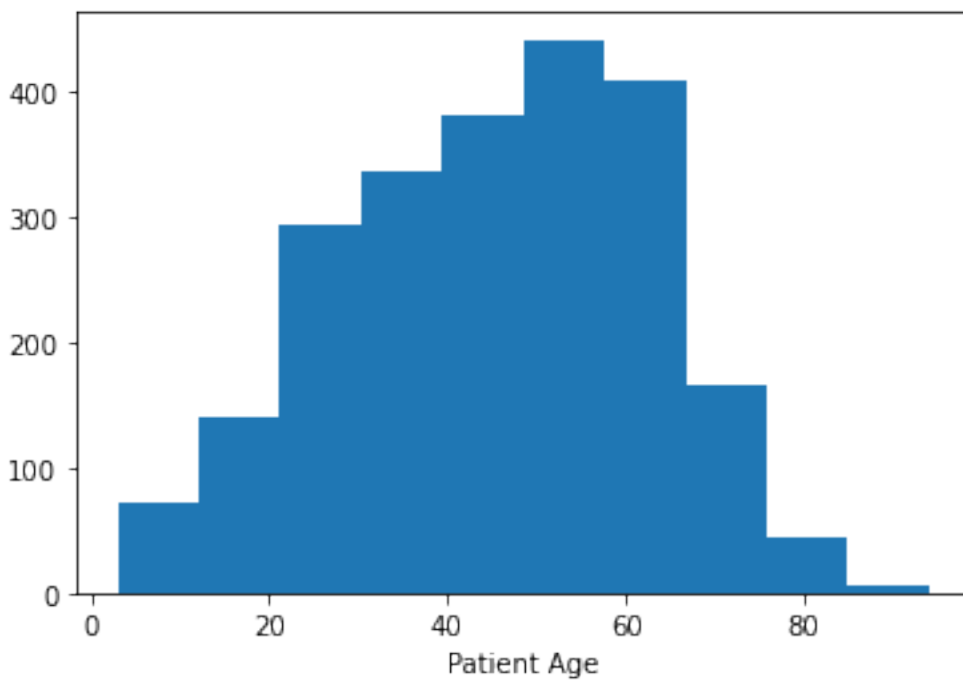
The threshold (0.20591968) was selected to achieve best F1 score of around 0.85 on validation set.

4. Databases

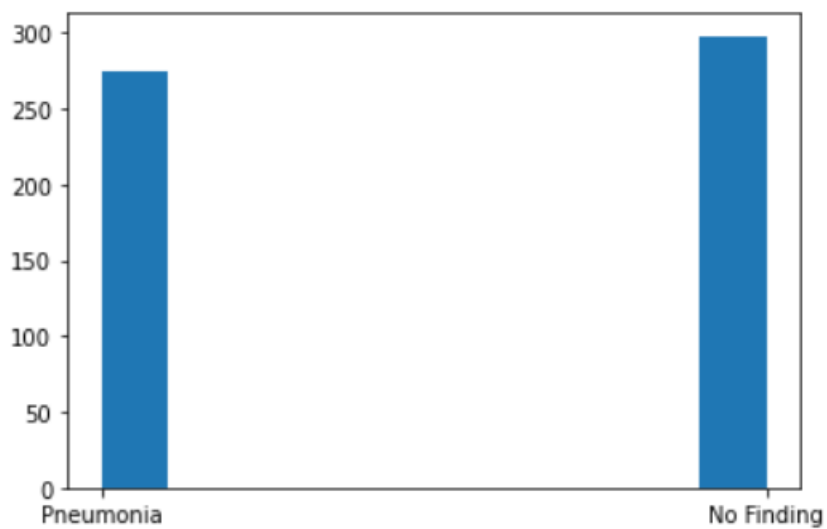
(For the below, include visualizations as they are useful and relevant)

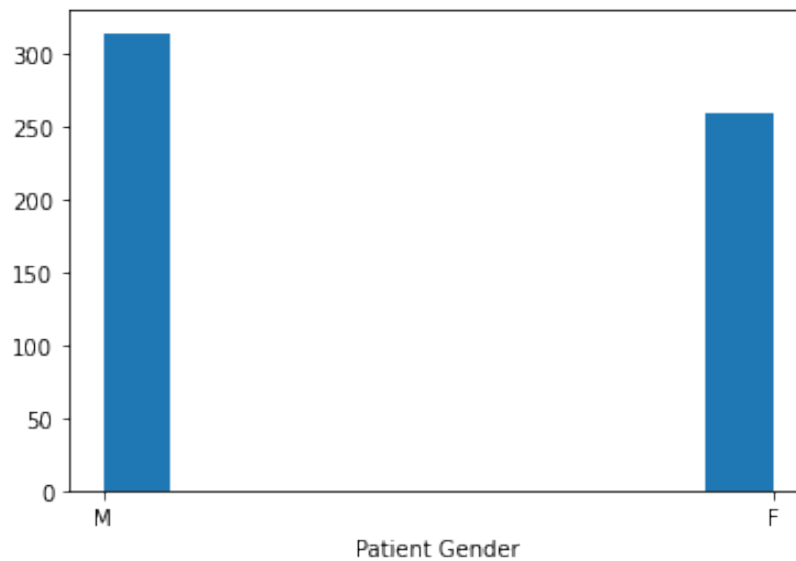
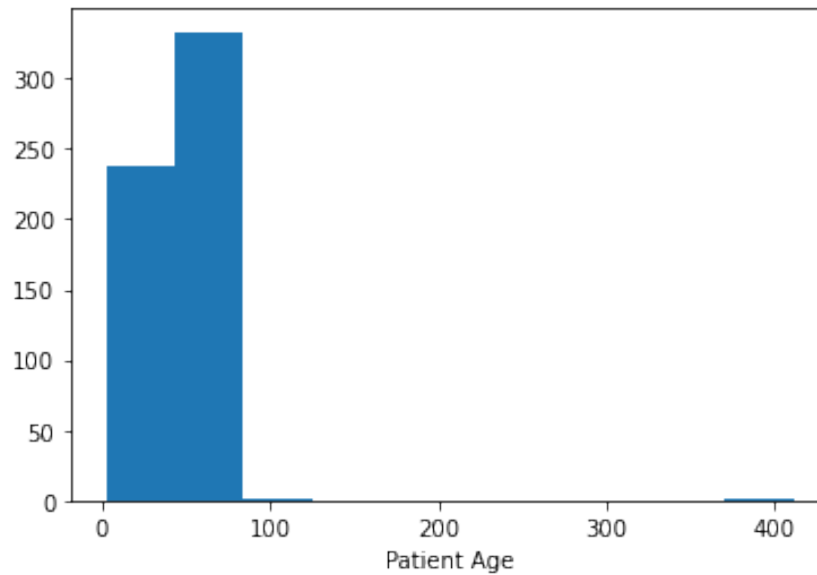
Description of Training Dataset:





Description of Validation Dataset:





5. Ground Truth

Natural Language Processing (NLP) to text-mine disease classifications from the associated radiological reports was used to generate ground truth. The disease labels are close to 90% accurate.

NLP is good approach to generate ground truth only if the accuracy of NLP model is very high (ideally 100%) otherwise Radiologist should be employed for generating ground truth.

6. FDA Validation Plan

Patient Population Description for FDA Validation Dataset

For FDA validation, data (total 500 patients same as validation dataset for algorithm building) should be generated for all ages between 0-100 with approximately equal number of males and females. The validation dataset should also have both xray images taken in both positions 'PA' and 'AP'.

Ground Truth Acquisition Methodology:

A well-trained Radiologist should examine xray images to determine the Ground truth.

Algorithm Performance Standard:

Algorithm should be judged on F1 score. F1 score should not less than 0.85.