FDA Submission

Your Name:Hadjer BENMEZIANE

Name of your Device: Pneumonia Detection Assistance on Chest X-Rays

Algorithm Description

1. General Information

Intended Use Statement: This software is intended to assist radiologist in detecting pneumonia on Chest X-Rays for women or men from the age of 20-100.

Indications for Use: The algorithm successfully detects pneumonia on chest x-rays when the patient is in AP or PA positions. It can be used for women or men from the age of 20-100.

Device Limitations: This algorithm must be run with a computer that meet minimum GPU and RAM requirements. The algorithm fails to detect pneumonia when presenting other chest diseases.

Clinical Impact of Performance: It is critical that the radiologist review all findings and proceeds to further analysis and tests to confirm the pneumonia.

2. Algorithm Design and Function
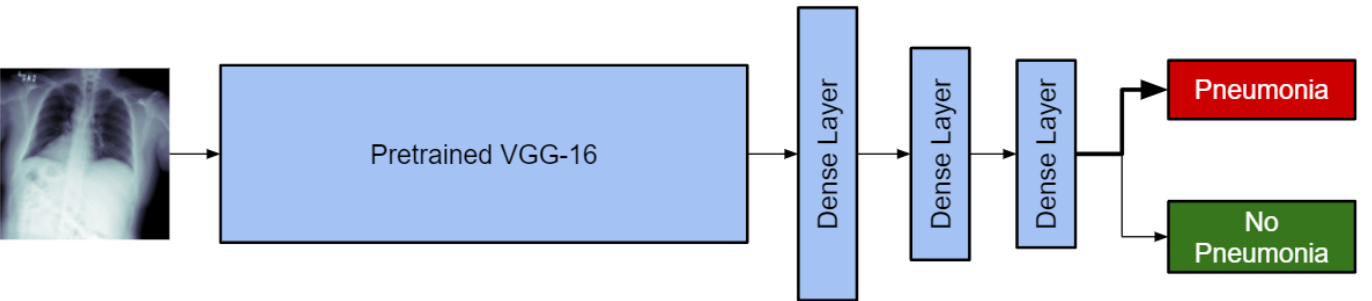
DICOM Checking Steps: Check if the DICOM is comform includes: Modality checking (DX), BodyPartExamined (CHEST), Patient position should be either "PA" or "AP".

Preprocessing Steps: Once the DICOM pixel array is opened, only normalization and recise (224x224) are applied.

CNN Architecture:

Layer (type) Output Shape Param
======= == VGG-16 (Model) (None, 7, 7, 512) 14714688

---

flatten_1 (Flatten) (None, 25088) 0

---

dense_1 (Dense) (None, 4096) 102764544

---

dropout_1 (Dropout) (None, 4096) 0

---

dense_2 (Dense) (None, 1024) 4195328

dropout_2 (Dropout) (None, 1024) 0

---

dense_3 (Dense) (None, 256) 262400

---

dense_4 (Dense) (None, 1) 257
====== === Total
params: 121,937,217 Trainable params: 107,222,529 Non-trainable params: 14,714,688

3. Algorithm Training

Parameters:

- Types of augmentation used during training:
    - custom horizontal blur
    - random zoom 0.1
    - horizontal_flip = True
    - height_shift_range = 0.1
    - width_shift_range = 0.1
    - rotation_range = 15
    - shear_range = 0.1
- Batch size: 32
- Optimizer learning rate: 1e-5
- Layers of pre-existing architecture that were frozen: all conv layers
- Layers of pre-existing architecture that were fine-tuned: final dense layers
- Layers added to pre-existing architecture: final dense layers
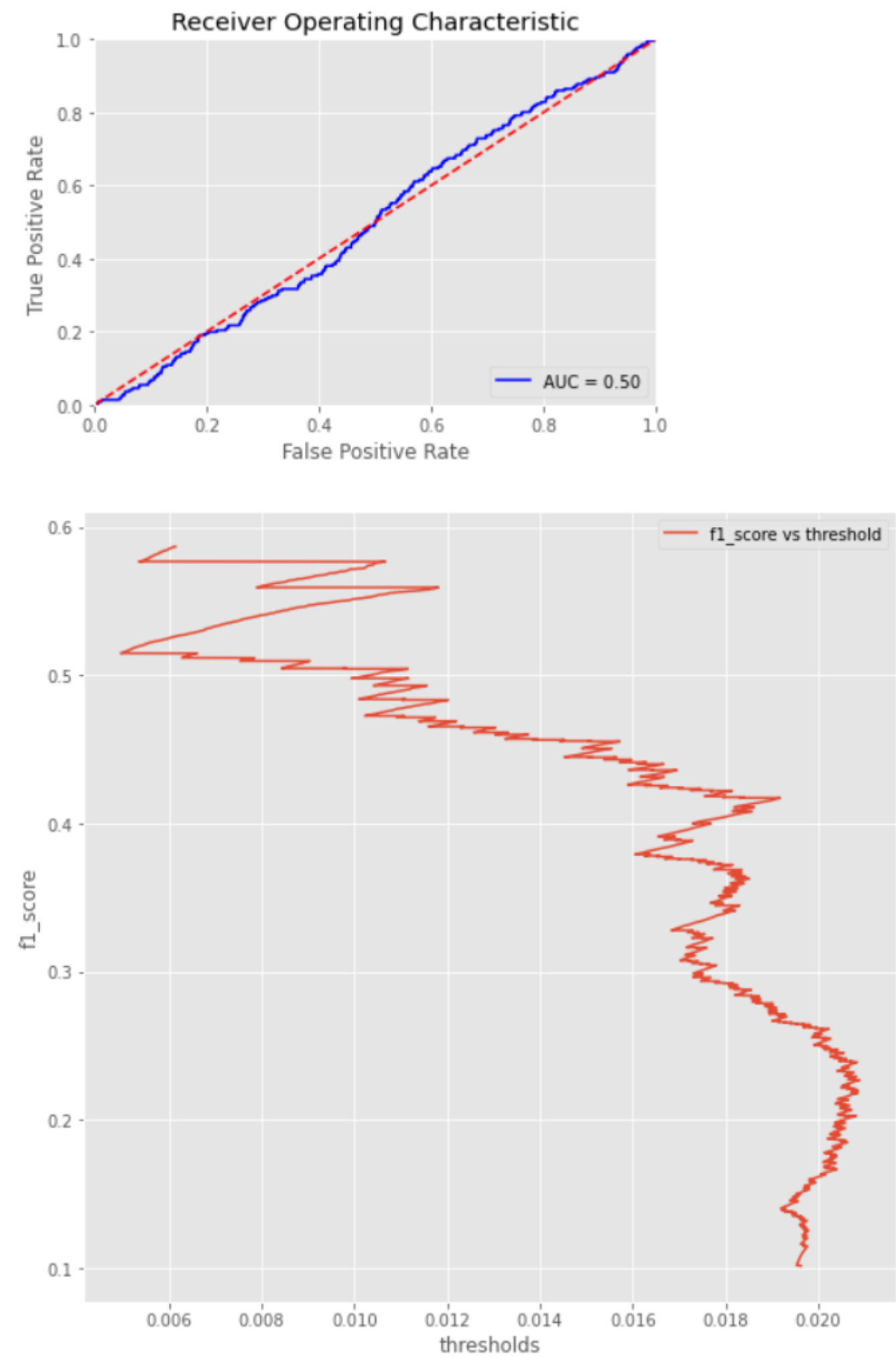
Final Threshold and Explanation: The performance of multiple thresholds [0, 1.0] with 0.01 step was explored by optimizing by ROC, by F1, and by maximizing Recall. Optimizing the f1-score is similar to optimizing ROC. Maximizing the recall to (1.0) gives an F1 accuracy of 0.019.

4. Databases

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

Description of Training Dataset: Our training set contains 2290 rows. Each row describe an image. The pneumonia_class is balanced with 50 percent of the data of patients with pneumonia. The gender in slightly imbalanced as we can see as follows, with more males than females that is similar to the global dataset.

Balance

Age
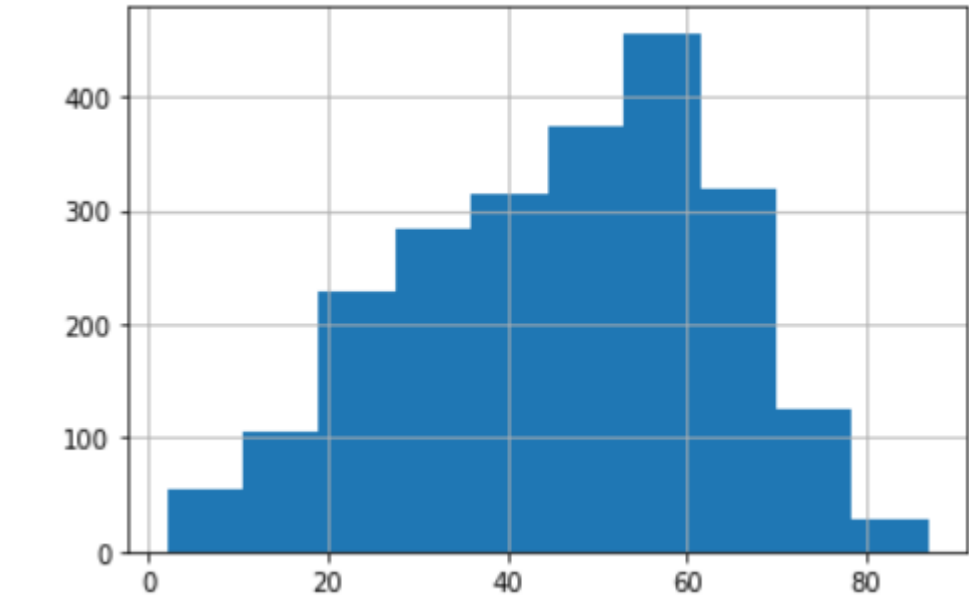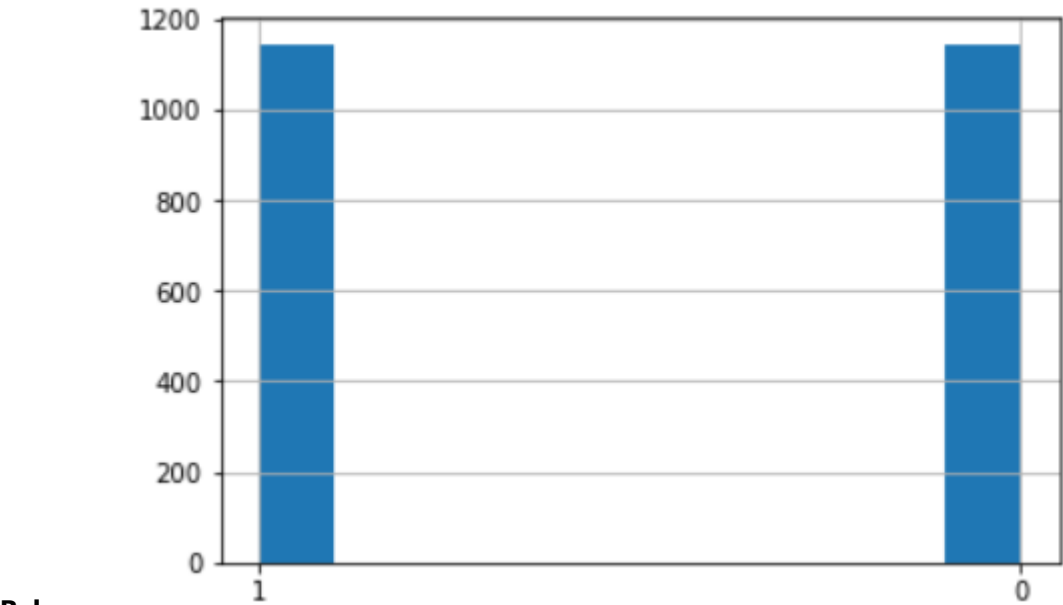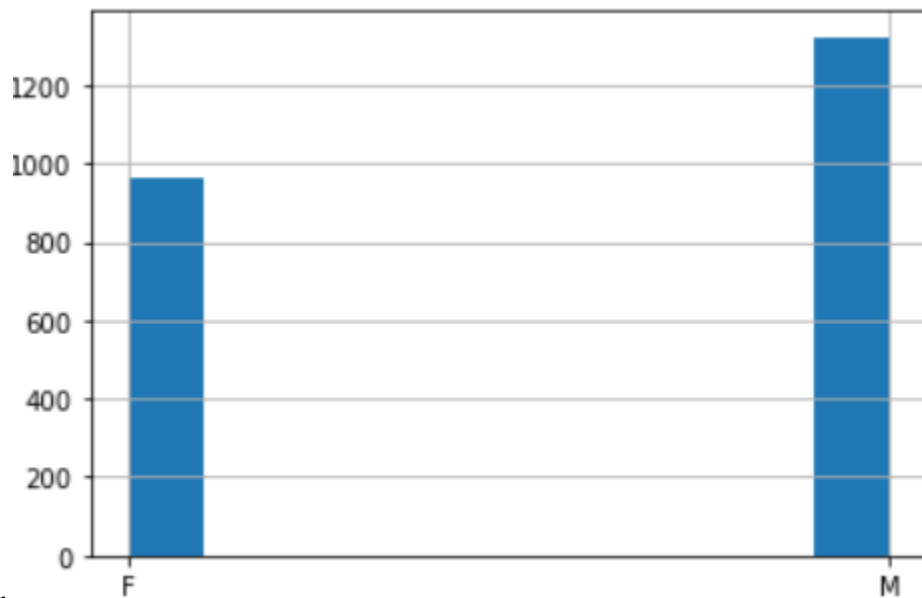
Gender

Description of Validation Dataset: The validation set contains 22359 rows. following the prevalence of pneumonia in the original dataset, we find 1% pneumonia cases in this set.
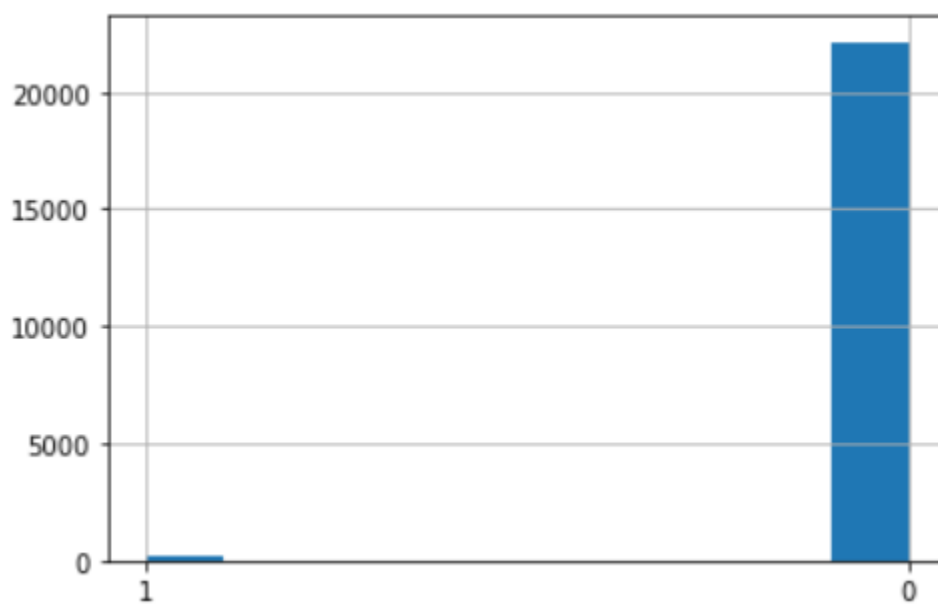
Balance

5. Ground Truth

NIH dataset was used for this project. It is comprised of 112,120 X-Ray images with disease labels from 30,805 unique patients. Each patients can have multiple diseases. This reflects real world scenarios. The labelling was done from radiologist reports which are expected to be >90% accurate.

6. FDA Validation Plan

Patient Population Description for FDA Validation Dataset: We need to make sure that the population used includes patient, female or men, from the age of 20-100, and with no previous pneumonia detected. The screening should be of "DX" modality and the only body part examined is the chest.

Ground Truth Acquisition Methodology: As the Ground Truth we will use the silver standard approach of using several radiologists since identifying Pneumonia is difficult for radiologists.

Algorithm Performance Standard: In this context, the recall is the most important metric we want to maximize. We select the model with recall = 1.0 The f1 score is 0.019 .