Proposal: Pneumonia or not?

The project will be based on the Kaggle project competition that started a couple of days ago and it will be running for the next two months.

This project will be split in three stages:
1. categorizing CXR between normal and abnormal
2. identifying pneumonia or not
3. identifying bounding boxes for areas of interest

The Data:

The dataset includes chest Xrays for patients for 25684 patients. The X-rays include normal X-rays, abnormal X-rays that do not have pneumonia, abnormal X-rays for patients with pneumonia.

The dataset includes a CSV file indicating for each patient if it was diagnosed with pneumonia and if the size of the bounding box or boxes for the areas of interest.

There are some images that seemed to be flipped, heart on the wrong side. Therefore, I will need to identify the flipped images and 'un-flip' them.

The images are in DICOM format and this have to be converted to a format that can be read by keras.

The images include age and gender. However, the data table does not have that information. Ideally, one could extract that information from the image and include it in the model. It seems that this information can be part of the DICOM standard tagging. Still, I need to investigate.

The models:

The main idea is to use transfer learning as used for this project with keras. The images are in DICOM format and this have to be converted to a format that can be read by keras.

Additional there is a large dataset of X-ray, CheXNet, (100,000) that could be used to train the model.

References:
Blog post by Roberto using transfer learning for diagnostics from Optical Coherence Tomographies https://www.robertoreif.com/blog/?offset=1530805351782

Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning, Daniel S. Kermany et al, RESOURCE| VOLUME 172, ISSUE 5, P1122-1131.E9, FEBRUARY 22, 2018

CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, Pranav Rajpurkar et al, https://arxiv.org/pdf/1711.05225v3.pdf, November 2017

https://towardsdatascience.com/detecting-pneumonia-with-deep-learning-studio-a1bd39ef1923