



Predicting Pneumonia

Addressing mortality rates in developing nations

July 2020 | Jonathan + Tom © ph[0]ton

Contents:

1. Context

2. Objective

3. Data

4. Base Model

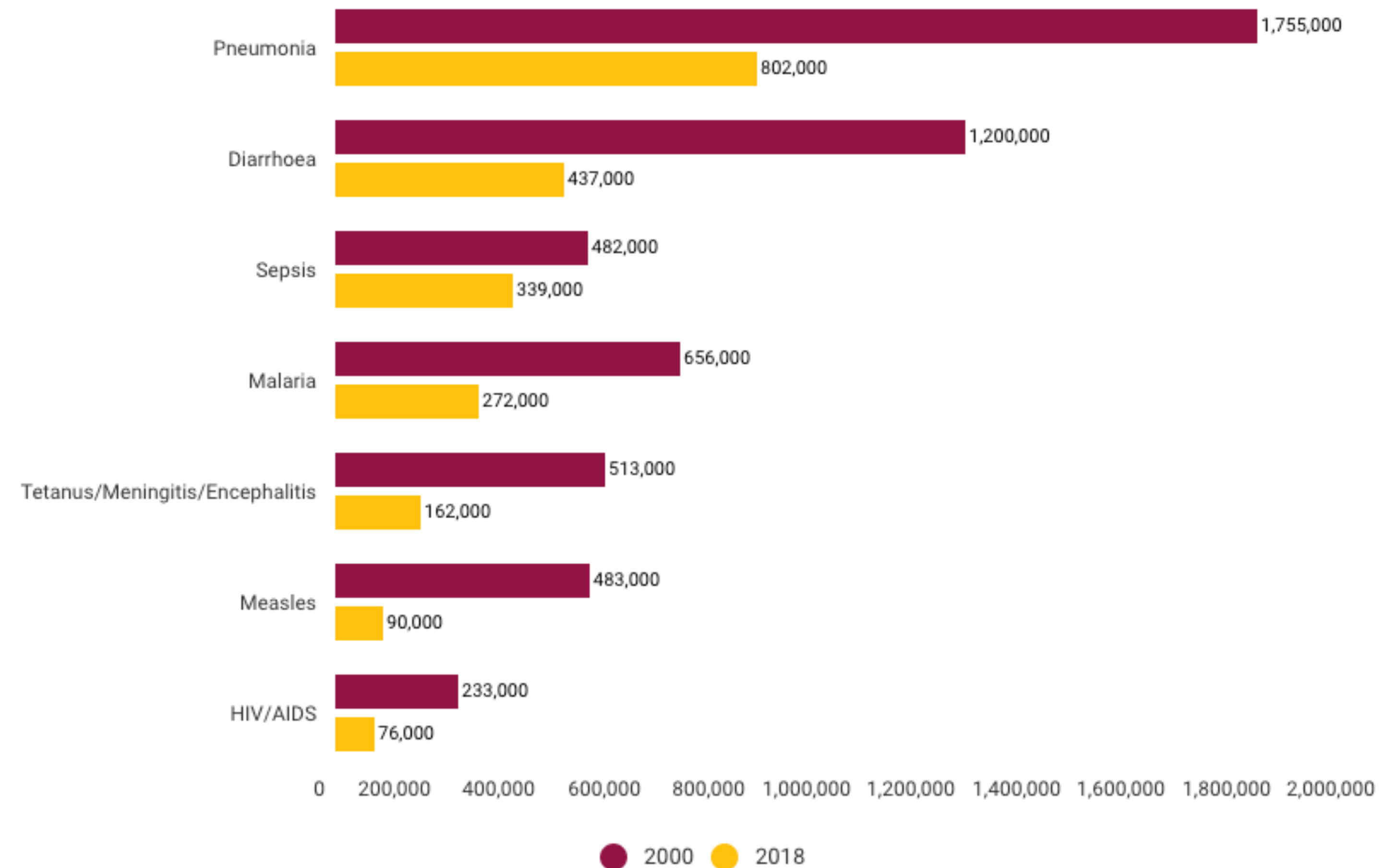
5. Transfer Learning

6. Next Steps

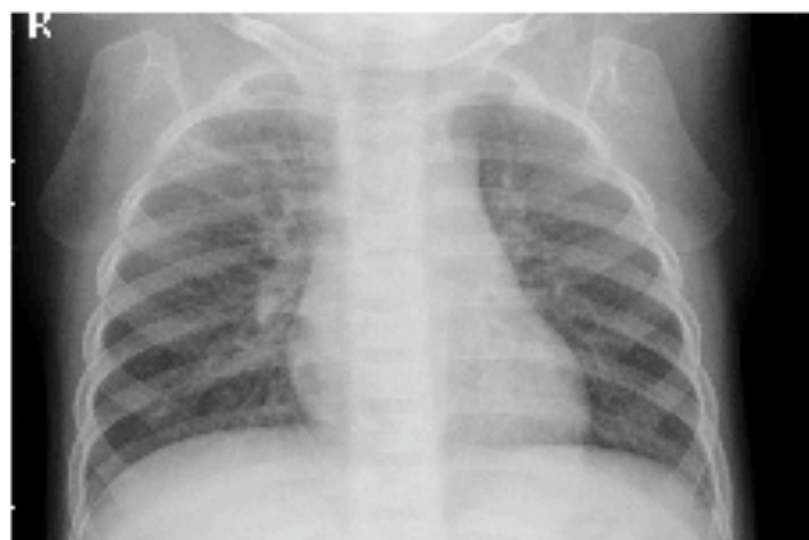
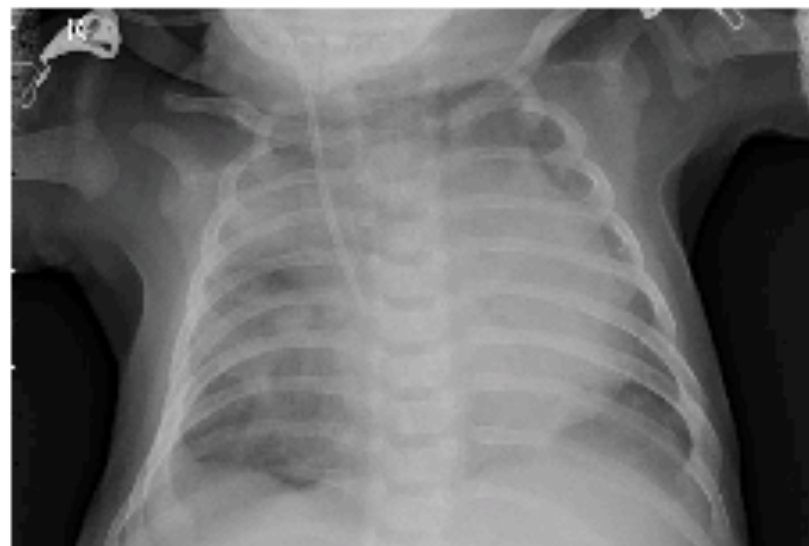
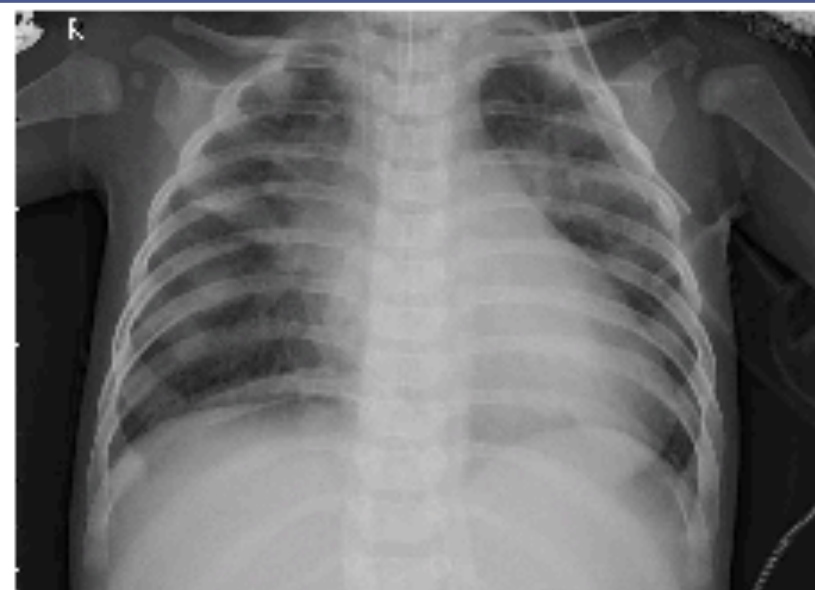
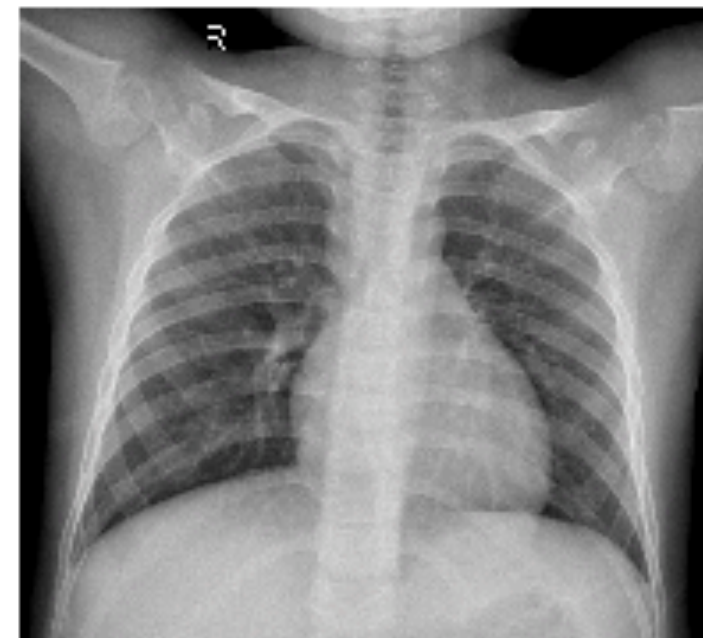
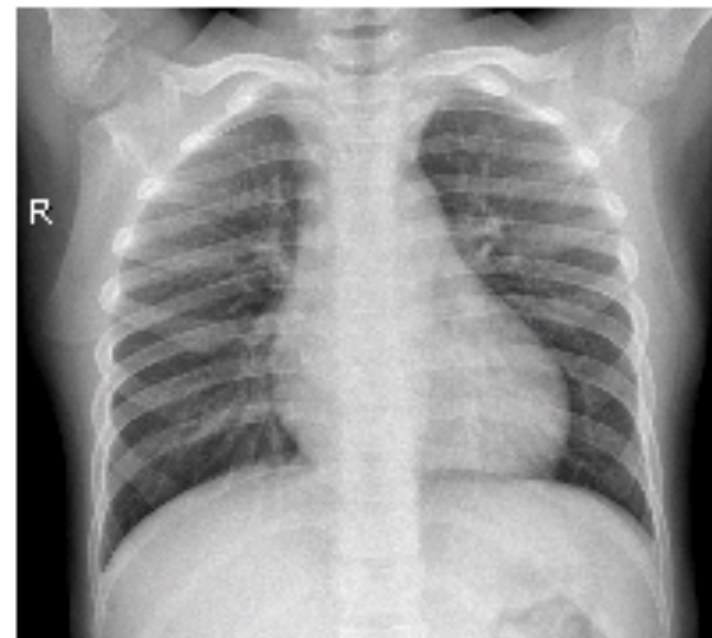
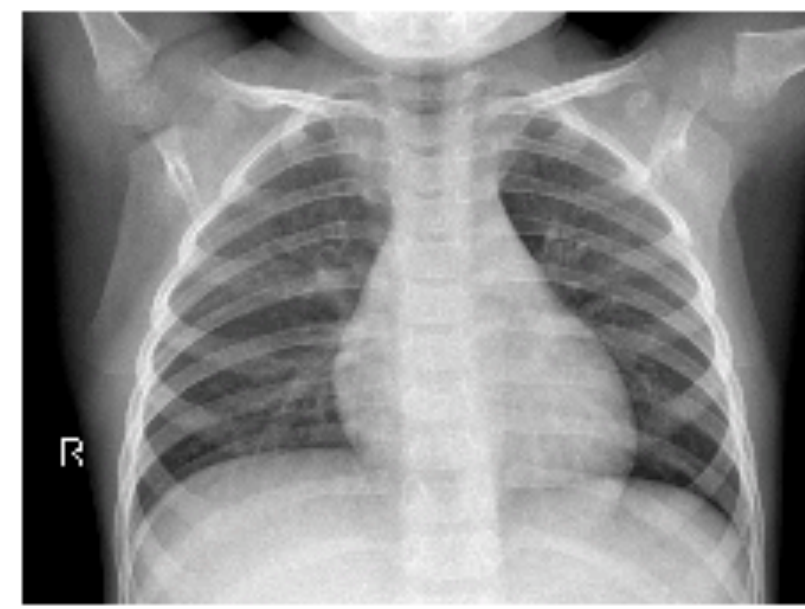
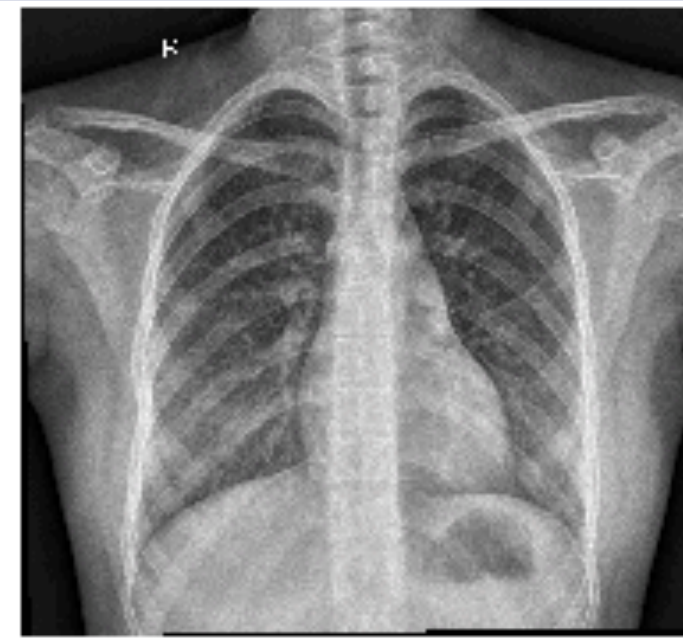
Context:

- Pneumonia kills more children than any other infectious disease, claiming the lives of over 800,000 children under five every year, or around 2,200 every day.
- Progress in reducing deaths due to pneumonia in children under five has been significantly slower than for other infectious diseases.
- Currently, 32% of children with suspected pneumonia are not taken to a health facility worldwide, while that number rises to 40 per cent for the poorest children in low- and middle-income countries.

Deaths of children under five by infectious disease, 2000 vs 2018



Source: UNICEF analysis based on WHO and Maternal and Child Epidemiology Estimation Group interim estimates produced in September 2019, applying cause of deaths for the year 2017 to United Nations Inter-agency Group for Child Mortality Estimation estimates for the year 2018.

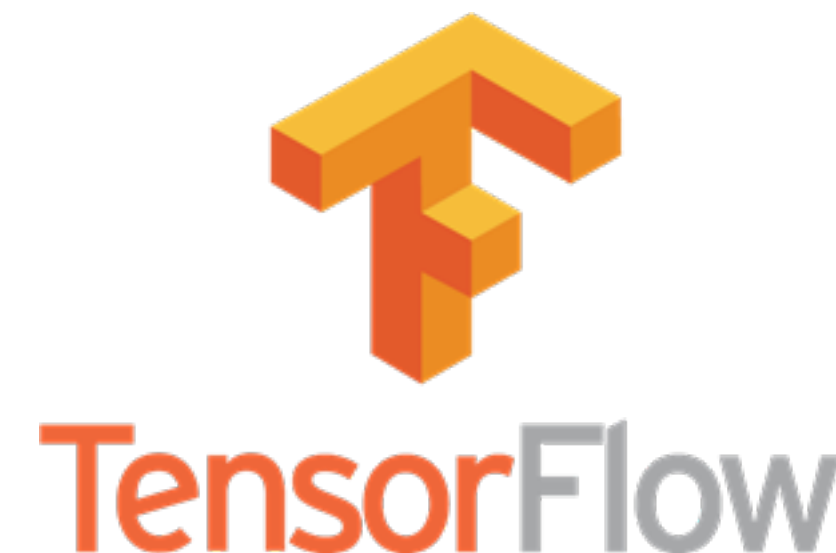
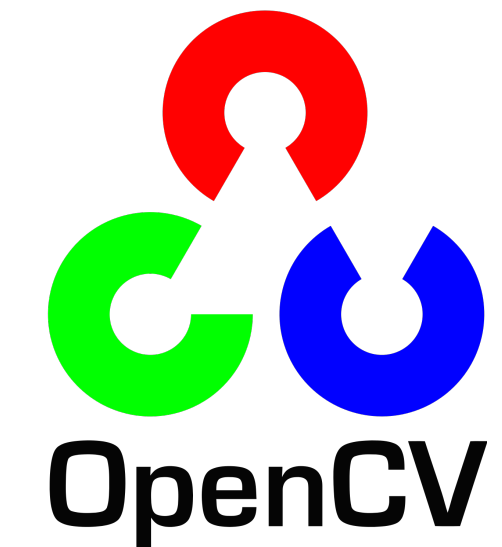


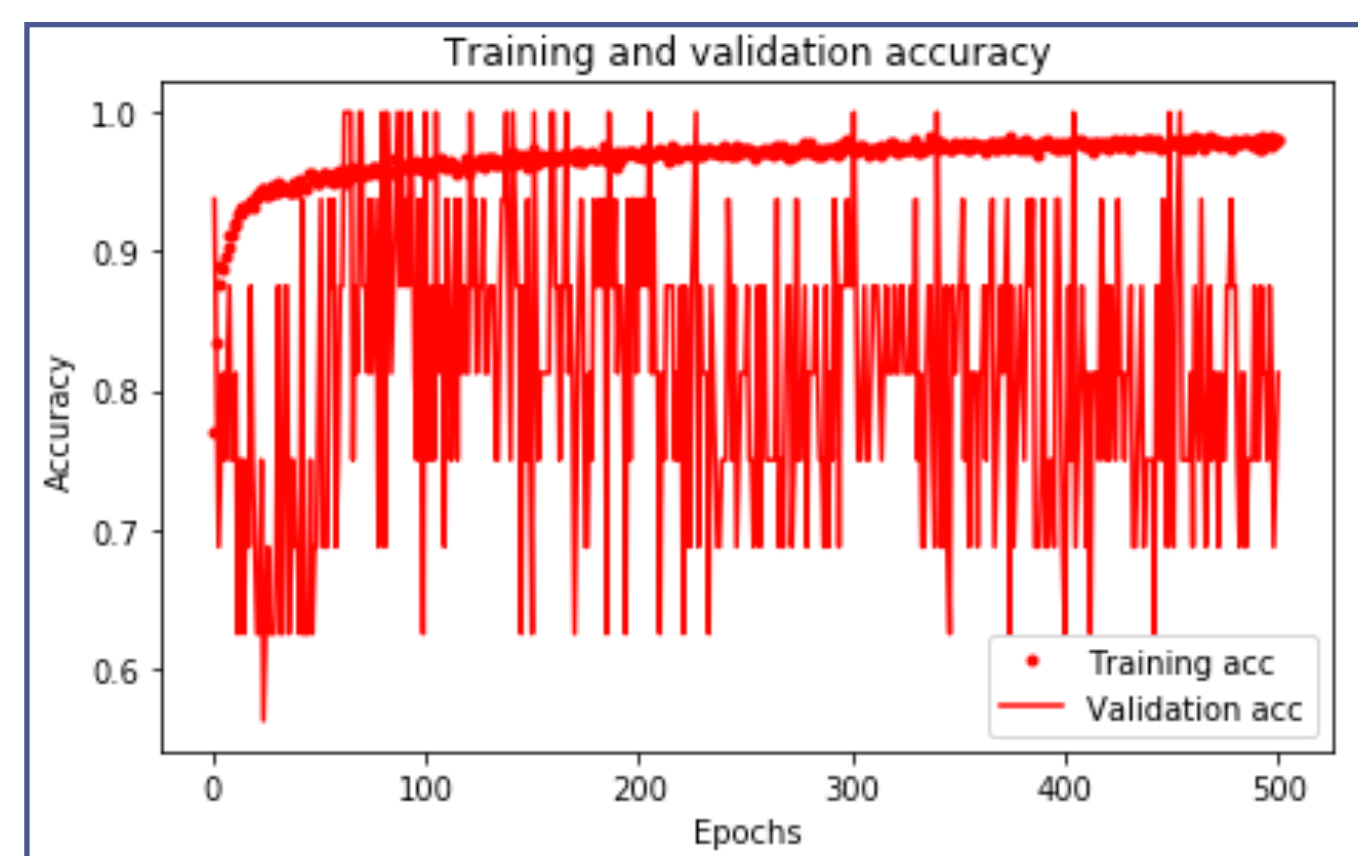
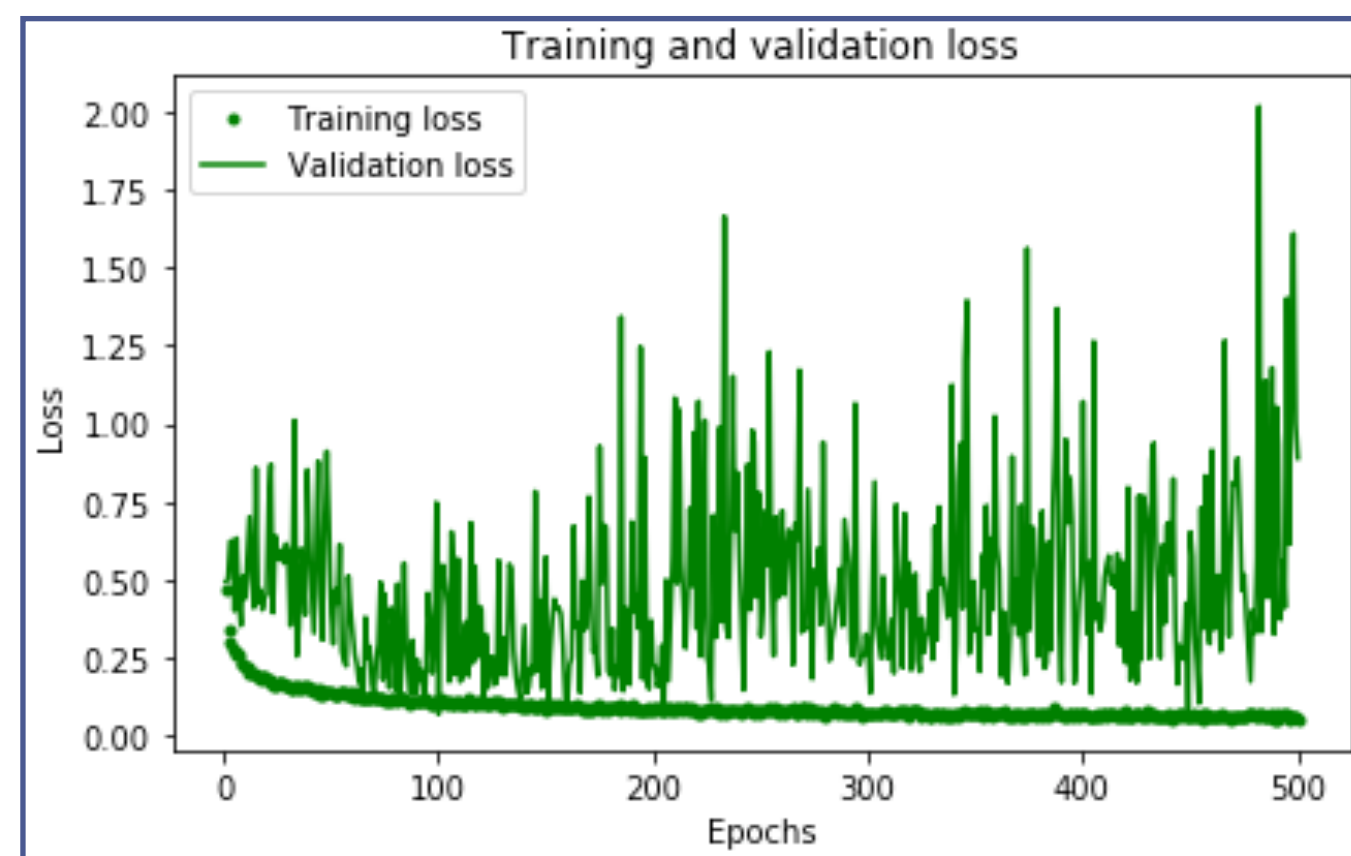
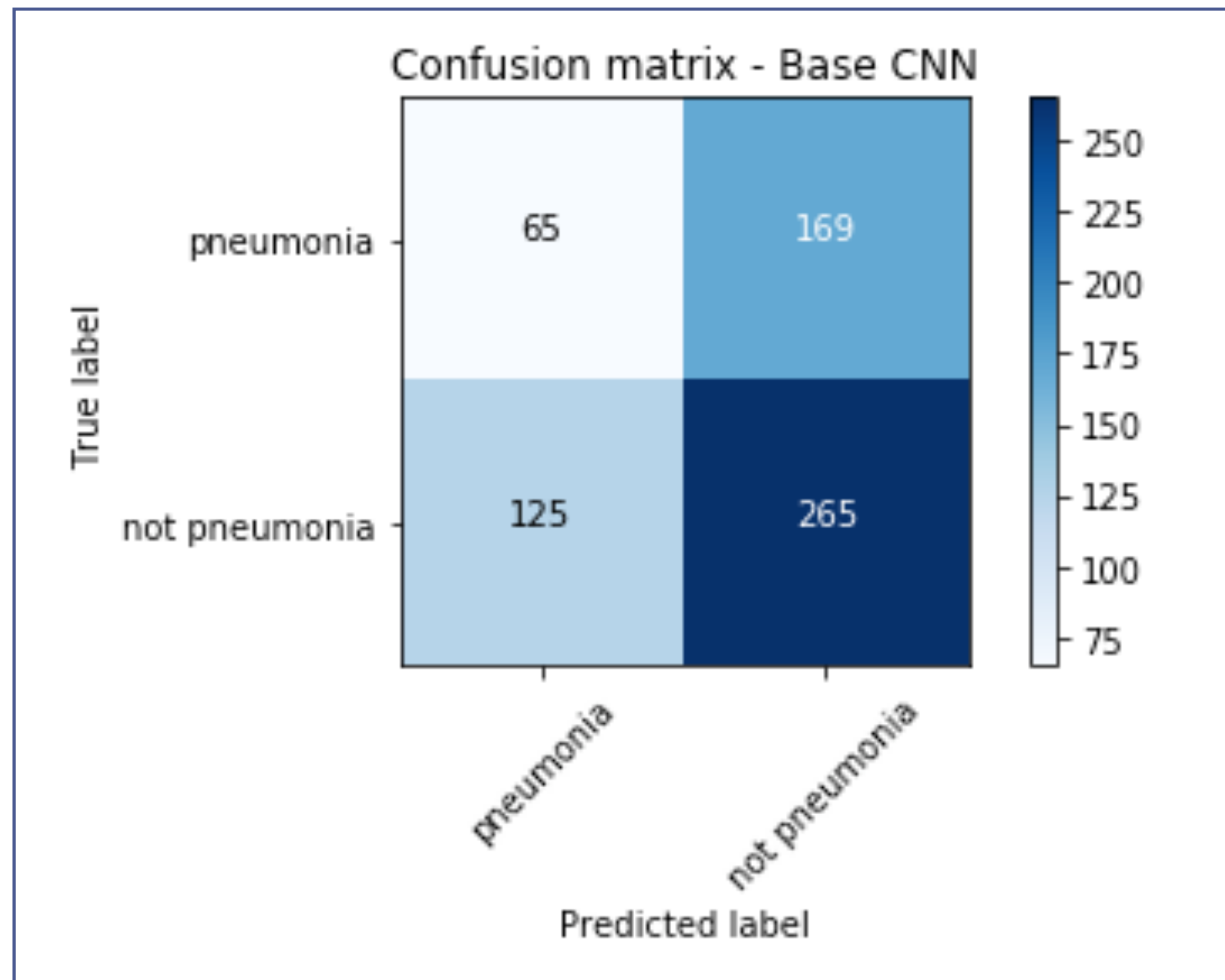
Objective:

- Worldwide, only about two-thirds of children receive the necessary help and care when pneumonia symptoms arise.
- As data from population-based survey data indicates, there has not been major progress in care-seeking behavior for pneumonia symptoms since 2000.
- A medical professional is needed to interpret X-rays, however there are not enough doctors in many developing countries.
- **Can we build a predictive model that will predict Pneumonia in pediatric chest X-rays?**
- Source: <https://data.unicef.org/topic/child-health/pneumonia/>

Data :

- All data for this investigation was provided by the Guangzhou Women and Children's Medical Center, Guangzhou.
- Download this dataset at: <https://www.kaggle.com/paultimothymooney/notebook>.
- The dataset is organized into 3 folders (train, test, val) and contains subfolders for each image category (Pneumonia/Normal).
- There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal).
- Libraries used for modeling:
 - Numpy
 - Pandas
 - Matplotlib
 - Bokeh
 - SKlearn
 - OpenCV
 - Tensorflow
 - Keras



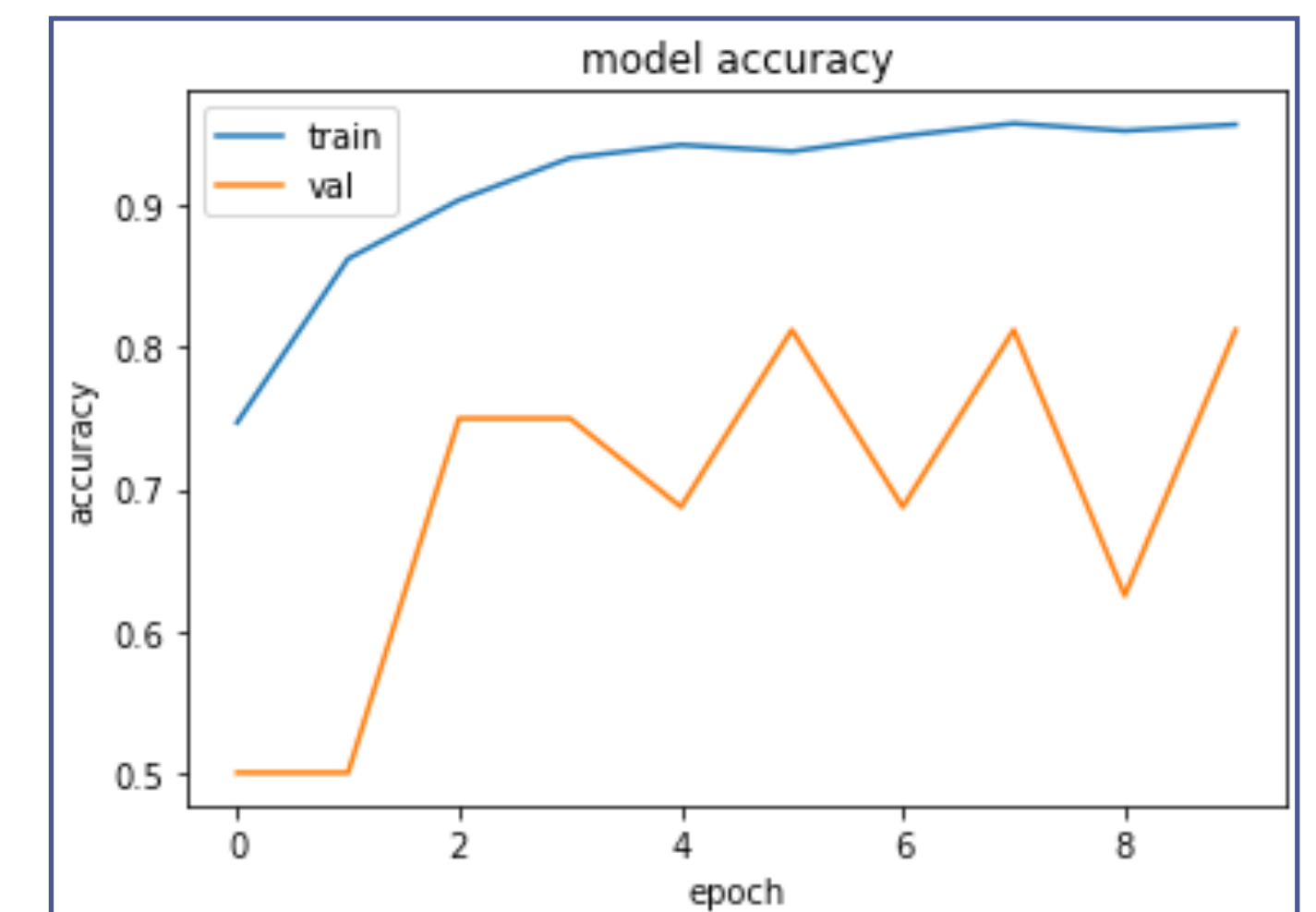
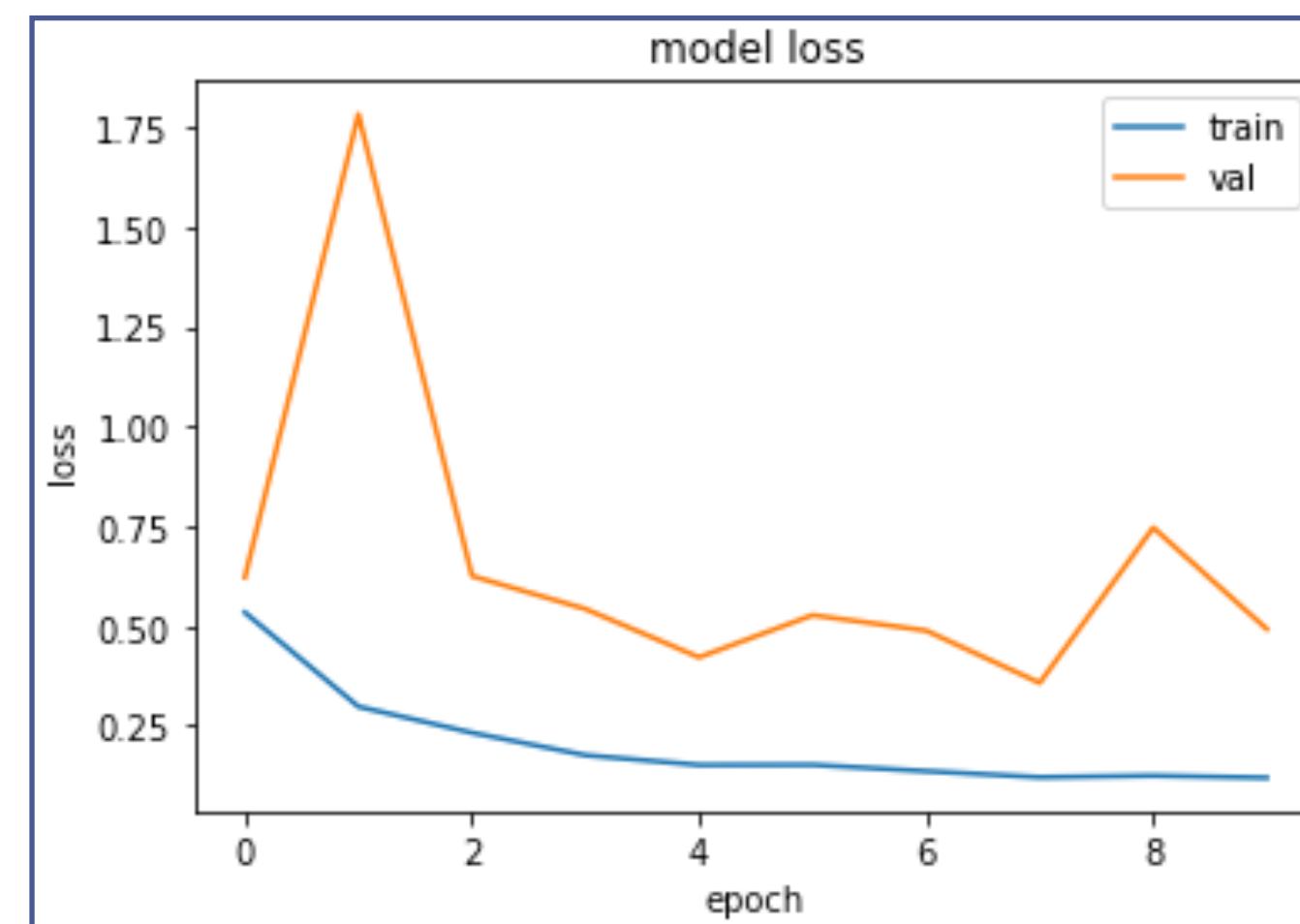
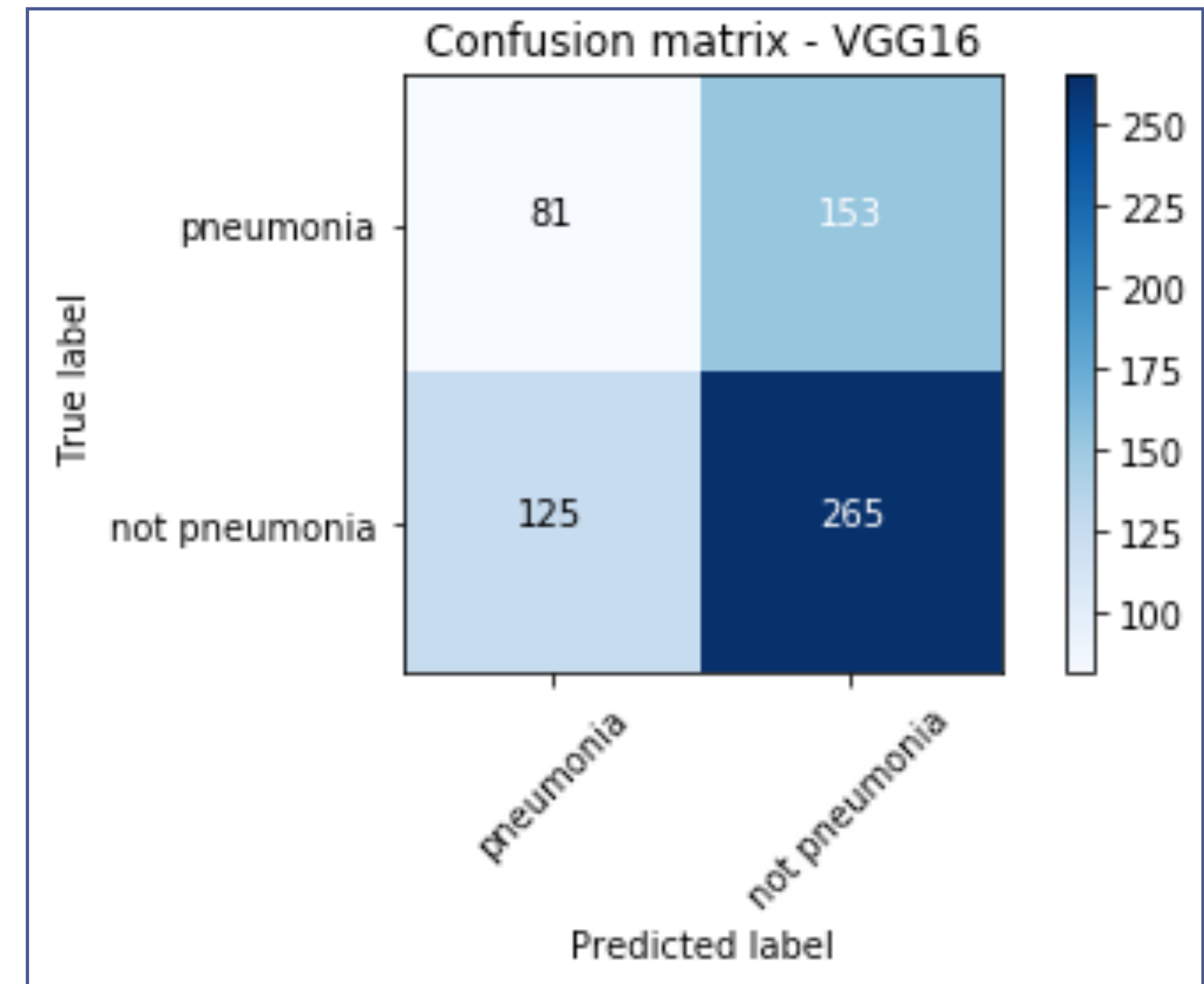


Base Model:

- For our first attempt we created a standard (vanilla) Convolutional Neural Network - CNN.
- It is made up of 2 Convolutional layers, 2 Pooling layers, Flatten, Dense, Dropout, and Output layer for a total of 221,025 parameters.
- Summary:
 - Resulting accuracy score after 500 epochs is 91%
 - Recall is the important metric here because we want to avoid False Negatives.
 - Recall score is 0.9872
 - However, the model is NOT robust enough given the high amount of False Negatives (169)

VGG16 Model:

- For our second attempt we used transfer learning to implement the VGG16 architecture.
- The model was created by K. Simonyan, A. Zisserman and consists of 10 convolutional layers, 10 zero padding layers, 10 max pooling layers, flatten, 2 dense, 2 dropout, and one output layer for 39,892,801 total parameters.
- Summary:
 - Resulting accuracy score after 10 epochs is 91%
 - Recall score is 0.9615
 - This model shows promise, in that it was better with predictions, however, many more epochs will need to be run in order to determine efficacy and potential for deployment.





**World Health
Organization**

ph[0]ton

Next Steps:

- We have an MVP.
- We need more data, more time for epochs, and to further treat the data to make our model even more accurate.
- Once complete, we would like to productize and partner with Unicef and W.H.O. to determine the best way to get this model in the hands of those who need it most.
- Finally, we would like to partner with American Hospitals to see if we can get older, portable, X-ray machines to ship our model with.



thank you

July 2020 | Jonathan + Tom © ph[0]ton