**Directory structure:**

* stage\_2\_train.csv - the training set. Contains patient Id’s and bounding box / target information.
* stage\_2\_sample\_submission.csv - a sample submission file in the correct format. Contains patient Id’s for the test set. Note that the sample submission contains one box per image, but there is no limit to the number of bounding boxes that can be assigned to a given image.
* stage\_2\_detailed\_class\_info.csv - provides detailed information about the type of positive or negative class for each image.

**My hardware:**

* Dell Inspiron Intel Core i7
* 16GB DDR4 RAM
* 1TB Hard Disk

**Software:**

* Windows 10
* Python 3.6.6
* Jupyter Notebook

**Data fields:**

* Patient Id \_- A patient Id. Each patient Id corresponds to a unique image.
* x\_ - the upper-left x coordinate of the bounding box.
* y\_ - the upper-left y coordinate of the bounding box.
* width\_ - the width of the bounding box.
* height\_ - the height of the bounding box.
* Target\_ - the binary Target, indicating whether this sample has evidence of pneumonia.

**Predictions:**

In this challenge we predict whether pneumonia exists in a given image. we did so by predicting bounding boxes around areas of the lung. Samples without bounding boxes were negative and contain no definitive evidence of pneumonia. Samples with bounding boxes indicate evidence of pneumonia.

When making predictions, we should predict as many bounding boxes as they feel were necessary, in the format: confidence x-min y-min width height

There should be only ONE predicted row per image. This row may include multiple bounding boxes.

A properly formatted row may look like any of the following.

For patient Id’s with no predicted pneumonia / bounding boxes: 0004cfab-14fd-4e49-80ba-63a80b6bddd6,

For patient Id’s with a single predicted bounding box: 0004cfab-14fd-4e49-80ba-63a80b6bddd6,0.5 0 0 100 100

For patient Id’s with multiple predicted bounding boxes: 0004cfab-14fd-4e49-80ba-63a80b6bddd6,0.5 0 0 100 100 0.5 0 0 100 100, etc.