Module 4 – Image Classification

Andy Peng

Hi my name is Andy Peng. Welcome to my presentation for Flatiron School Module 4 – Image Classification. For the extent of this project, our stakeholders are a company that focuses on predicting whether a patient has Pneumonia or not given patient's chest x-ray image to lessen the workload of doctors.

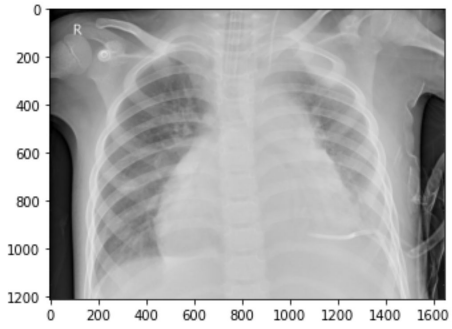
Overview

- Images
- Modeling

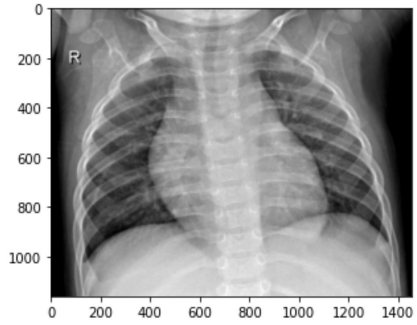
We managed to create a model that can predict whether the patient have Pneumonia with decent accuracy while also minimizing the chances of classifying a patient as false negative. Before we talk about the results of our model, we will be showing a few images to show different features our model is looking for when training.

Chest X-Ray Images

<matplotlib.image.AxesImage at 0x7fcd7677b748>



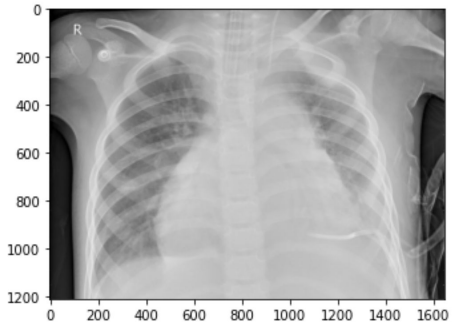
<matplotlib.image.AxesImage at 0x7fcdec38f390>



Given these two chest x-ray images of two patients, are you able to classify which patient has Pneumonia?

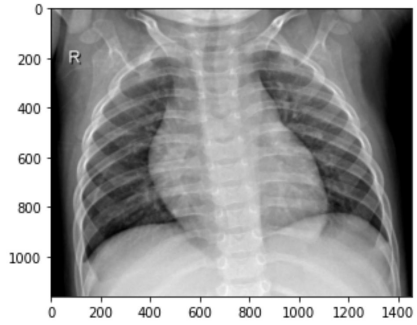
Chest X-Ray Images

<matplotlib.image.AxesImage at 0x7fcd7677b748>



Pneumonia

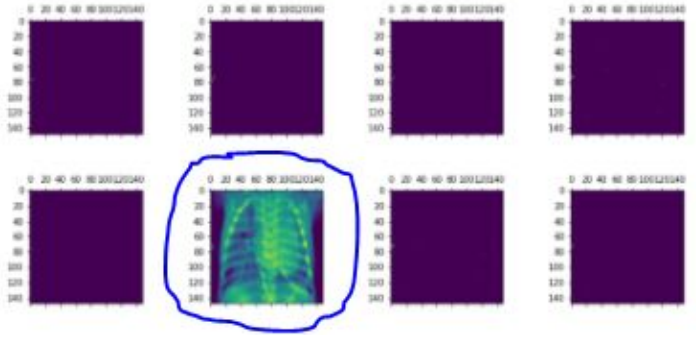
<matplotlib.image.AxesImage at 0x7fcdec38f390>



Normal

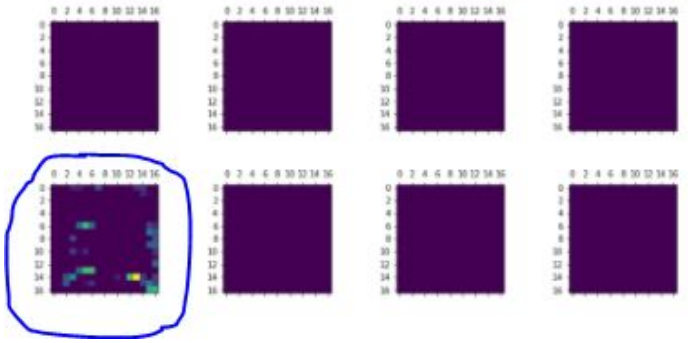
If you guessed the left image, then you are correct. But what certain features in the image do the computers look for when identifying whether a patient has Pneumonia or not? Let's take a look at some in on our convolutional neural network.

Channels of first activation layer



In the first layer of our convolution neural networks, here are the channels of the first convolutional layer. Notice in this image that we are basically looking at the outline of the chest x ray image. Pay attention to how detail the outline of the image is.

Channels of sixth activation layer



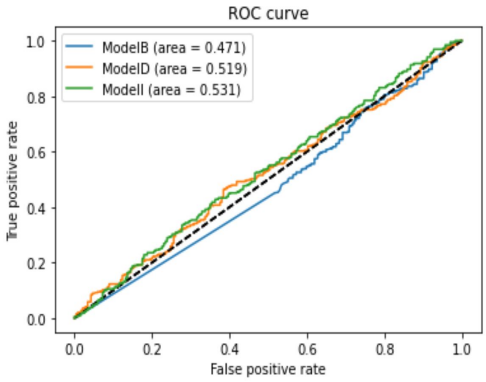
Now take a look at the channel images in our sixth layer of our model. As you can see as we move down our model we will look at simpler details in our image.

Modeling

- Convolution Neural Network
  - Baseline Model - Convolution Neural Network
  - Regularization Applied
  - Imbalance issues fixed

For modeling we used convolution neural networks. As a base model, we just built a general convolution neural network. The second model was built upon our baseline model by adding in dropout layers, regularizers and lower learning rates. And finally the last model included all that we discuss before and fixing the imbalance issues in our data. In our data set, pneumonia images were three times as many as normal images. After training these three models, here are the results.

Results



- ModelI - AUC

Here is a ROC curve displaying the AUC results for the different models we did. For best AUC score, it would be our ModelI. Although our model is doing slightly better than randomly guessing, we can further improve our model which will be discuss at the end.

Results

	Model	Recall Score	Precision Score	F1 Score	Testing Accuracy	AUC Score
0	modelB	0.599359	0.534569	0.524400	0.599359	0.501709
1	modelD	0.552885	0.526310	0.534024	0.552885	0.495299
2	modelI	0.570513	0.552769	0.558420	0.570513	0.521368

- Accuracy/ Recall - Baseline Model
- Precision/F1 Score - Imbalance Model

Here are the results of our three models. For best accuracy and recall score we would use our baseline model, ModelB. For best precision and F1 score we would use our Imbalance model, ModelI.

Recommendations

	Model	Recall Score	Precision Score	F1 Score	Testing Accuracy	AUC Score
0	modelB	0.599359	0.534569	0.524400	0.599359	0.501709
1	modelD	0.552885	0.526310	0.534024	0.552885	0.495299
2	modelI	0.570513	0.552769	0.558420	0.570513	0.521368

Since we are working to classify patients on whether they have Pnemuonia or not given a patient's chest x ray. Our goal would be to minimize the amount of patients we classify as normal when they indeed do have Pneumonia because we wouldn't want a patient to get sent home when they are sick. In other words we want to maximize our recall score. If you pay attention to the table, modelB's recall score and testing accuracy is slightly better than modelI by 0.02, but modelI is better than modelB in precision, F1 and AUC scores. Therefore, we would be sticking with modelI.

Next Steps

- Doctor's Method
- Cross Validation
- Gather More Data

For our next steps we can try to learn what features doctors look at in a chest xray to determine whether a patient has Pneumonia or not. Not only can we ask for a doctor's opinion, we could also use cross validation or gather more data to further improve our model.

Thank You