# Machine learning project info for PHYS 555

Mukesh Aryal

03/31/2022

## Project description and objectives

Machine learning is getting very popular in the medical field. Deep learning techniques to correctly identify and segregate abnormalities on medical images such as X-rays, MRIs and CT scans are gaining more and more attention. The deep learning algorithms are not only used for detection of the diseases but are also being incorporated in the development of drugs and vaccines.

The dataset I am considering for applying deep learning is a set of medical x-ray dataset that I also used for the 1st part of the course. The dataset is available in *Kaggle*. The dataset includes x-ray images of people with no Pneumonia and people who has Viral and Bacterial Pneumonia. The normal chest X-ray portrays clear lungs without any areas of abnormal opacification in the image. The bacterial pneumonia includes a so-called focal lobar consolidation whereas viral pneumonia shows more diffuse "interstitial" pattern in both lungs. There are 5,863 X-Ray images in JPEG format. All chest radiographs are claimed to have been screened for quality control by eliminating low quality or unreadable scans.

Outline of the project:

1. *Data Pre-Processing*: The images are of different lengths and widths thus all images would be resized to, for example, 220x220x3. The greyscale images would then be converted to 3 channel and the image pixels would be normalized.

2. *Building CNN*: The next step would be to make the convolutional neural network utilizing the libraries. The network will be used to correctly classify and identify whether the images pertain to normal patients or people with different types of Pneumonia. The problem would be a classfication problem with a maximum of 3 classes.

3. *Transfer learning*: I also plan to use transfer learning to observe the benefits of using pre-trained neural networks. I am considering Alexnet, Inception net or VGG16 to apply transfer learning.

4. *Evaluation, comparison, and conclusion*: Since I used the same dataset earlier in the 1st part of the course, I am planning to compare the final results from the deep learning part with the one I got previously. I will also compare the result that I will get from my own implementation of CNN with that of transfer learning.