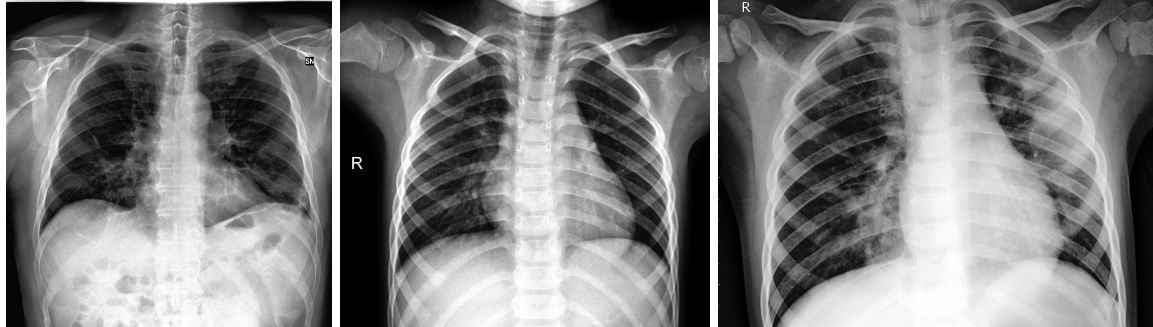


Lung X-Rays: Can We Accurately Detect Viruses?

DS 4002 Case Study - Pedro Del Valle



Prompt: The 2020 Pandemic has struck! In its wake, we as a society have become increasingly aware of global health concerns and our individual contributions to the spread (or lack thereof) of contagious diseases. Unfortunately, it can be quite difficult to assess whether somebody has a viral disease such as COVID-19 or other infections. Many individuals tend to display what we now call “asymptomatic” behavior, in which they are infected and can transmit the virus without having any reason to suspect they are ill. The problem is further magnified within the healthcare industry. That’s where you come in. Doctors at UVA have been taking X-Ray images of several patients, some of whom are healthy, some of whom have Coronavirus (COVID-19), and some of whom have Viral Pneumonia. However, these doctors cannot perfectly establish a means to assess someone’s state of health by looking at these X-Rays. The images are too similar to each other to pinpoint differences with the human eye. For this reason, the medical staff needs to find some automated method by which they can predict whether a patient has healthy lungs, Coronavirus, or Pneumonia based on their lung X-Rays. You have been chosen to aid in this task.

Deliverable: Your job for this high-stakes project is to build a machine learning algorithm that can view this image data and predict the corresponding patient’s condition with an 80% effective rate. For this task, you have been given two sets of data. The train set of data is meant to feed your model images of healthy patients, Coronavirus patients, and Pneumonia patients so that it can learn from them and establish patterns to distinguish between the three. You will then use your constructed model on the test dataset to assess how accurate the predictions are. Attached to this document, you will find the rubric and additional materials.