**Patient information columns (Don’t use these in model):**

* ImagePath
  + Relative path to the respective x-ray images which will be input to the model.
* PatientID
  + Unique ID for each patient.
  + Use this to group each patient and avoid leaking same patient scans into train/test data.
* Date
  + Scan Date
* Time
  + Scan Time

**Demographic information columns (Can be as auxiliary input/features for model):**

* Gender
* Age

**Clinical information columns (Can be as input/features for model):**

* Comorbidity
  + Indicates if the patient has any pre-existing condition during the x-ray.
  + Positive status indicated by - “C”
* Category
  + The radiologist has indicated a preliminary categorization of the x-ray into
    - P (positive) – abnormalities visible in x-ray
    - N (negative) – abnormalities not visible in x-ray
  + Please note that positive/negative status only indicates how severe the x-ray looks and **not covid-19 status**. All scans are covid-19 positive
* OldScoring
  + Some older scans had a slightly different scoring scale by the radiologist. So, I have added this column so that model can adjust accordingly. You can choose to use this as a feature if necessary.
    - True – indicates older scoring scheme
    - False – indicates latest scoring scheme
* SpO2/time
  + Blood oxygen saturation level.
  + Chest abnormalities and SpO2 score might be correlated.
  + It is recommended to use this as input feature

**Chest X-ray scores (CXR Score - Target output/labels for model):**

The radiologist has scored each x-ray according to severity of visible abnormalities on x-ray. He has scored each section of lung separately. We will use broadly 2 sections (Left & Right lung) **OR** a combined total score.

* Lt Total
  + Chest X-ray score (CXR) for left lung
* Rt Total
  + Chest X-ray score (CXR) for right lung
* Total Score
  + Lt Total + Rt Total

These scores can be **categorized and used instead of scores**, if you want to use a classifier instead of regression. The thresholds for classification are as follows

CXR < 12 – Mild

12 < CXR < 20 – Moderate

20 – Severe

Note that these scores are developed by the radiologist and classification thresholds are recommended based on their observation. It might change in future as more data is analyzed.