**COVID-19 Radiography Dataset Description**

**Dataset Link:** [data-soure](https://www.kaggle.com/datasets/preetviradiya/covid19-radiography-dataset%20)

**Overview:**

The COVID-19 Radiography Dataset is designed for research and development in medical imaging, focusing on the detection of lung conditions such as COVID-19 and other forms of lung opacity. This dataset provides a diverse collection of chest X-ray images to support the training and evaluation of deep learning models aimed at improving diagnostic accuracy.

**Components:**

**1. Image Data:**

- **Type:** Chest X-Ray Images

- **Format:** JPEG

- **Resolution:** Up to 224x224 pixels for model training

- **Total Count:** 16,204 images

**2. Data Split:**

**- Training Set:**

- Purpose: To train machine learning models to classify and recognize chest X-ray images.

- Composition: Includes a balanced mix of images labeled as Lung Opacity and Normal.

- Preparation: Images are preprocessed, resized, and divided into training and validation subsets. The exact number of images per category should be specified based on dataset proportions.

**- Validation Set:**

- Purpose: To fine-tune model parameters and validate model performance during training.

- Composition: A balanced mix of Lung Opacity and Normal images.

- Preparation: Used to monitor and adjust the model’s performance, ensuring that the data reflects the diversity of the training set.

**- Test Set:**

- Purpose: To evaluate model performance on unseen data, providing an unbiased assessment of generalization.

- Composition: Includes both Lung Opacity and Normal images that were not part of the training or validation sets.

- Preparation: Used for final evaluation metrics such as accuracy, precision, recall, and F1-score.

**3. Source and Usage:**

- **Source:** Compiled from publicly available X-ray images and medical imaging repositories.

- **Usage:** Intended for developing and assessing machine learning models for detecting lung opacity and differentiating it from normal lung conditions in chest X-rays.

**Summary:**

The COVID-19 Radiography Dataset provides a robust foundation for developing diagnostic models by offering a substantial number of annotated X-ray images. The careful splitting into training, validation, and test sets ensures a balanced representation of Lung Opacity and Normal conditions, facilitating effective model training and evaluation