

Data Collection and Preprocessing Phase

Date	4 July 2024
Team ID	SWTID1720097611
Project Title	Covid Vision: Advanced COVID-19 Detection From Lung X-Rays With Deep Learning
Maximum Marks	2 Marks

Data Quality Report Template

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Inconsistent Image Sizes	Low	Using ImageDataGenerator for Preprocessing, setting target size while loading the image dataset. This ensures all the images loaded are in same size
Kaggle Dataset	Imbalanced Data Distribution	Moderate	Using Under-Sampling to Reduce the number of images in Class having more images