Data Collection and Preprocessing Phase

Date	2 Dec 2024
Team ID	TMID739650
Project Title	ADVANCED COVID-19 DETECTION USING LUNG X-RAYS BY 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
Dataset	Missing values in critical columns	High	Impute missing values using median or mean imputation
Dataset	Duplicate records in the dataset	Moderate	Remove duplicates based on unique identifiers
Dataset	Incorrect data types (e.g., numerical stored as text)	High	Convert columns to correct data types (e.g., string to float)
Dataset	Missing labels for classification task	High	Create a labeling system or use a proxy for missing labels