Data Collection and Preprocessing Phase

Date	18 March 2024
Team ID	SWTID1720165000
Project Title	CovidVision: Advanced COVID-19 Detection From Lung X-Rays With Deep Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	The primary objective of this project is to develop a robust and efficient machine learning model that can accurately detect COVID-19 from lung X-ray images. The model aims to distinguish between four categories of lung conditions: normal, lung opacity, COVID-19, and viral pneumonia.
Data Collection Plan	Kaggle was used for the dataset
Raw Data Sources Identified	Covid-19 Radiography Database: Author: Dr. Muhammad Chowdhury, Amith Khandakar, Tawsifur Rahman.

	It consists of Lung x-ray images of Normal, Lung opacity, Viral pneumonia and Covid.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	It consists of Lung x-ray images of Normal, Lung opacity, Viral pneumonia and Covid.	https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database	images	777MB	Public