Data Collection and Preprocessing Phase

Date	29 June 2024
Team ID	SWTID1720084679
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	This project aims to develop a deep learning model to detect COVID-19 from lung X-ray images. The model will be trained and validated using a comprehensive dataset from Kaggle, with the ultimate goal of aiding healthcare professionals in the timely and accurate diagnosis of COVID-19.
Data Collection Plan	The data will be collected from the Kaggle dataset "COVID-19 Radiography Database".
Raw Data Sources Identified	This dataset consists of X-ray images of patients diagnosed with COVID-19, normal cases, viral pneumonia, and lung opacity. It provides a balanced set of images to train and evaluate the deep learning model for detecting COVID-19.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
COVID-19 Radiography Database	A comprehensive collection of X-ray images including COVID-19 positive cases, normal cases, viral pneumonia cases, and lung opacity cases. This dataset provides a balanced set of images necessary for training, validating, and testing the deep learning model for COVID-19 detection.	kaggle datasets download -d tawsifurrahman/covid19-radiography-database	Image	816 MB	Public