Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWTID1720440447
Project Title	Covid Vision: 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 goal is to develop an advanced, precise, and real-time COVID-19 detection tool utilizing lung X-rays through transfer learning. This tool is intended to aid radiologists and healthcare professionals in swiftly and effectively diagnosing COVID-19, thereby enhancing patient outcomes and bolstering public health efforts..
Data Collection Plan	Data obtained from Kaggle is downloaded to Google Drive and accessed via Google Colab using the opendatasets Python library. Two directories are created: one for images labeled as 'Covid' and another for images labeled as 'Normal'.
Raw Data Sources Identified	The lung X-ray dataset is sourced from Kaggle.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Covid -19 Detection from Lung X-rays	The dataset contains data of chest X-ray images for Covid-19 positive cases along with Normal and Viral Pneumonia images.	https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database	Image	816 MB	Public