Data Collection and Preprocessing Phase

Date	09-07-2024
Team ID	SWTID1720433291
Project Title	CovidVision: Advanced COVID-19 Detection From Lung X-Rays With Deep Learning
Maximum Marks	2 Marks

Data Collection Plan :

Section	Description
Project Overview	The goal of this project is to develop an advanced AI system that utilizes deep learning techniques to rapidly and precisely detect COVID-19 from lung X-ray images. The aim is to enhance diagnostic capabilities, alleviate healthcare system burdens, and improve patient outcomes by providing a reliable and scalable diagnostic tool.
Data Collection Plan	1. Identify datasets related to COVID-19 detection that leverage lung X-rays and deep learning techniques. 2. Prioritize datasets that have a balanced distribution of COVID-19 positive, pneumonia, and normal cases. This will help ensure robust model training and validation.
Raw Data Sources Identified	The raw data sources for this project is taken from kaggle website. The provided sample data contains a subset of information such as covid-19 status x ray images etc for analysis.

Raw Data Sources report :

Source Name	Description	Location/URL	Format	Size	Access Permissions

Covid-19 Radiography Database	Collection of chest X-ray images for covid detection and related research	https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database	Images and CSVs	1.16 GB	Public, requires kaggle account
CoronaHack chest x-ray Dataset	A dataset of chest X-ray images for diagnosing COVID-19 and other related illnesses	https://www.kaggle.com/datasets/praveengovi/coronahack-chest-xraydataset	Images And CSVs	1 GB	Public,Requires a Kaggle account