

## Data Collection and Preprocessing Phase

Date	15 March 2024
Team ID	SWTID1720439521
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	This Deep Learning model aims to speed up and accurate the process covid diagnosis. Using a dataset that contains X-ray images of normal, people suffering with covid and people suffering with pneumonia, this model seamlessly detects and diagnosis the person using his chest X-ray.
Data Collection Plan	Search for datasets containing the X-ray images of the mentioned types.
Raw Data Sources Identified	The raw data collection sources of this model involves kaggle and youtube

### Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
kaggle	This dataset consists of 3 folders containing 3 types of x-ray images. This dataset needs to be split manually	<a href="https://www.kaggle.com/code/rollanmaratov/covid19-detection-using-tensorflow-from-chest-xray/data">https://www.kaggle.com/code/rollanmaratov/covid19-detection-using-tensorflow-from-chest-xray/data</a>	image	3 GB	Public
Youtube	This dataset is pre-split into test and train. Therefore you just need to add variables into the model and start with pre-process	<a href="https://drive.google.com/drive/folders/1Ax8PtO8sC2_uWFwqo570lz-4Q3gCsu9">https://drive.google.com/drive/folders/1Ax8PtO8sC2_uWFwqo570lz-4Q3gCsu9</a>	Image	1 GB	public