**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 01 May 2025 |
| Team ID | 739942 |
| Project Title | CovidVision: Advanced COVID-19 Detection From Lung X-Rays With Deep Learning Using IBM Cloud |
| 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 | To develop deep learning model capable of accurately detecting COVID-19 infections from lung X-ray images, techniques, hosted on IBM Cloud. CovidVision will streamline the diagnosis process, reduce the burden on medical staff, and contribute to early detection and timely treatment of COVID-19 cases. |
| Data Collection Plan | **Data will be sourced from publicly available dataset of chest X-ray images labeled with COVID-19 and other relevant conditions by metadata for model training.** |
| Raw Data Sources Identified | The raw dataset identified is consist as images of chest X-ray from COVID-19, bacterial pneumonia, and non-COVID pneumonia. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle covid19 detection using Tensorflow from chest Xray | A dataset containing labeled COVID-19 cases infections from lung X-ray images (e.g.,bacterial pneumonia, and non-COVID pneumonia) | https://www.kaggle.com/code/rollanmaratov/covid19-detection-using-tensorflow-from-chest-xray/data | Image | 2GB | Public |
|  |  |  |  |  |  |
|  |  |  |  |  |  |