**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | 739877 |
| Project Title | WCE Curated Colon Disease Classification using Deep Learning |
| Maximum Marks | 2 Marks |

**Data Collection Plan & Raw Data Sources Identification**

**Data Collection Plan**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | Develop a deep learning model to classify colon diseases from WCE and colonoscopy images, aiding faster diagnosis and better treatment outcomes. |
| Data Collection Plan | Data will be sourced from publicly available colonoscopy and WCE image datasets with disease labels, supplemented by patient metadata for model training. |

**Raw Data Sources**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle WCE Curated Colon Disease Dataset | A dataset containing labeled colonoscopy images (e.g., ulceratives, esophagatis polyps).. | [**https://www.kaggle.com/datasets/francismon/curated-colon-dataset-for-**](https://www.kaggle.com/datasets/francismon/curated-colon-dataset-for-deep-learning)[**deep-learning**](https://www.kaggle.com/datasets/francismon/curated-colon-dataset-for-deep-learning) | Image | 2 GB | Public |