**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 03 July 2024 |
| Team ID | SWTID1720085445 |
| Project Name | Hydration Essentials: Classifying Water Bottle Images |
| 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 project "Hydration Essentials: Classification of Water Bottle Levels" aims to develop a deep learning model that can classify the level of water in a bottle using images. This application could be useful in various scenarios, such as monitoring hydration levels in healthcare settings, smart home devices, or fitness applications. |
| Data Collection Plan |  **Image Acquisition:** Gather a diverse set of images of water bottles with varying water levels.   **Labeling:** Manually label the images with corresponding water levels (e.g., empty, low, half, full).   **Data Augmentation:** Apply techniques such as rotation, scaling, and flipping to increase the diversity and robustness of the dataset.   **Normalization:** Scale pixel values to a standard range, |
| Raw Data Sources Identified | <https://www.kaggle.com/datasets/chethuhn/water-bottle-dataset> |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Water-bottle-dataset | This dataset contains three folders namely over-flown, full water , half water.  Totally it consists of 486 images. | <https://www.kaggle.com/datasets/chethuhn/water-bottle-dataset> | Image | 67MB | Private (with access) |