

Data Collection and Preprocessing Phase

| | |
|---------------|--|
| Date | 20 July 2024 |
| Team ID | SWTID1720082030 |
| Project Title | Hydration Essentials: Classifying Water Bottle |
| 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 | Hydration Essentials: Classifying Water Bottle Images"" is a project that uses computer vision and machine learning techniques to classify images of water bottles into different categories based on their features and characteristics. The project aims to streamline and automate the process of categorizing water bottles for e-commerce platforms, inventory management, and quality control purposes. By training a model on a diverse dataset of water bottle images, including variations in size, shape, material, and color, the project can quickly and accurately classify new images, providing valuable insights for manufacturers, retailers, and consumers. |
| Data Collection Plan | Search for datasets with water bottle images with different levels of water. |

| | |
|-----------------------------|--|
| Raw Data Sources Identified | Collected datasets from different open sources like kaggle.com, data.gov, UCI machine learning repository etc. The dataset used for this project was obtained from Kaggle. |
|-----------------------------|--|

Raw Data Sources Template

| Source Name | Description | Location/URL | Format | Size | Access Permissions |
|-------------|---|---|--------|-------|--------------------|
| Dataset 1 | The dataset comprises of different water bottle images(Full water level, Half water level,Overflowing). | https://www.kaggle.com/datasets/chethuhn/water-bottle-dataset | CSV | 67 MB | Public |