**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 22 JUNE 2024 |
| Team ID | 739950 |
| Project Title | Anemiasense: Leveraging Machine Learning For Precise Anemia Recognitions |
| Maximum Marks | 2 Marks |

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

This template provides a structured framework for planning and documenting the data collection process and identifying raw data sources for the "Anemiasense" project. Customize the template based on the specific sources and methods relevant to your project's needs and objectives..

**Data Collection Plan Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | Provide a brief overview of the "Anemiasense" project, highlighting its goals and objectives related to leveraging machine learning for precise anemia recognition. |
| Data Collection Plan | **Source**: Extract from database in the form of CSV files. |
| Raw Data Sources Identified | Hospital Database: Anonymized patient records containing hematological profiles and demographic information.  National Health Surveys: Surveys conducted periodically to assess anemia prevalence, nutritional habits, and health-related behaviors among different population segments. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Dataset 1 | Description of the data in this source. | anemia.csv - Google Drive..  <https://drive.google.com/file/d/1KMJFNFGwoaQoAouIPabMEHcT1bvqEXau/view?usp=sharing> | CSV | 35KB | Public |