**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 July 2024 |
| Team ID | xxxxxx |
| Project Title | Early Prediction of Chronic Kidney Disease |
| Maximum Marks | 2 Marks |

**Data Quality Report Template**

The Data Quality Report Template will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Source** | **Data Quality Issue** | **Severity** | **Resolution Plan** |
| Dataset 1 | **Missing Values:** **Blank**  **Inaccuracies:** **Incorrect**  **Inconsistencies:** **Conflicting**  **Outliers:** **Extreme**  **Duplicates:** **Repeated**  **Errors:** **Mistake**  **Timeliness:** **Outdated**  **Bias:** **Skewed** | Moderate | **Missing Values:** **Impute Fillna method**  **Inaccuracies:** **Validate**  **Inconsistencies:** **Standardize**  **Outliers:** **Filter**  **Duplicates:** **Remove**  **Errors:** **Correct**  **Timeliness:** **Update**  **Bias:** **Adjust** |
| Dataset 1 | Categorial data in the dataset | moderate | Encoding has to be done in the data. |