**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 2024 |
| Team ID | 740061 |
| Project Title | Early Prediction Of Chronic Kidney Disease |
| 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 | This project aims to predict the CKD in early stages and give better treatment to them patients. By using Machine Learning we can predict the CKD. |
| Data Collection Plan | Search for data in different sources like kaggle.com, data.gov, UCI machine learning repository etc. |
| Raw Data Sources Identified | The raw dataset identified is UCI machine learning repository. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Dataset | We can get all information of patients (age, sugar, rbc, wbc, hemoglobin etc.,) | <https://drive.google.com/file/d/1mPl4yaTKuKZ3017YfYC19Ni7Y964eCNI/view?usp=sharing> | CSV | 48Kb | Public |