Data Collection and Preprocessing Phase

Date	3th July 2024
Team ID	SWTID1720090815
Project Title	Early Prediction Of Chronic Kidney Disease Using Machine Learning
Maximum Marks	2 Marks

Data Quality Report Template

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

Data Source	Data Quality Issue	Severity	Resolution Plan
UC Irvine – ML Repository	The dataset had several missing values, along with additional columns which needed to be removed. The missing values needed to be handled.	High	Handling missing values through statistical filling i.e. mean and mode. Columns removed on the basis of multicollinearity and through Recursive Feature Elimination (RFE).