Data Collection and Preprocessing Phase

Date	23 september 2024
Team ID	LTVIP2024TMID25021
Project Title	Prediction and Analysis of liver patient data using machine learning
Maximum Marks	2 Marks

Data Quality Report:

The Data Quality Report 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 Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Missing Values Albumin and globium ratio	Moderate	Use mode imputation
Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data.