

Data Collection and Preprocessing Phase

Date	03-10-2024
Team ID	LTVIP2024TMID24892
Project Title	Liver Patient Identification – prediction of liver patient
Maximum Marks	2 Marks

Data Quality Report

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
Kaggle Dataset	Missing values are in the ['Albumin_and_Globulin_Ratio']	Low	Used mean/median imputation.
Kaggle Dataset	Categorical data in the dataset	Moderate	There is no need of Standardization and Normalization for the dataset, as we have used Ensemble Technique.