Data Collection and Preprocessing Phase

Date	28th June 2025
Team ID	LTVIP2025TMID41511
Project Title	Revolutionizing Liver Care: Predicting Liver Cirrhosis Using Advanced Machine Learning Techniques.
Maximum Marks	2 Marks

Data Quality Report Template

This report summarizes the data quality issues identified in the liver cirrhosis dataset, along with their severity levels and proposed resolution plans. The goal is to systematically identify and rectify discrepancies to ensure high-quality data for accurate predictions.

Data Source	Data Quality Issue	Severity	Resolution Plan
Kaggle Dataset	Missing values in all the columns of the dataset. (42 columns)	High	Use mean/median imputation.

Kaggle Dataset	Categorical data in the dataset	Moderate	Perform encoding (e.g., Label Encoding or One-Hot Encoding).