Data Collection and Preprocessing Phase

Date	16 july 2024
Team ID	SWTID1720074204
Project Title	prediction and analysis of liver patient data using ml
Maximum Marks	2 Marks

Data Quality Report Template

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
Dataset	Albumin_and_Globulin_Ratio' feature contain 4 NaN values.	Moderate	import pandas as pd df['Albumin_and_Globulin_Ratio'] = df['Albumin_and_Globulin_Ratio'].fillna(df['Albumin_and_Globulin_Ratio'].median())
…	…	…	…