Data Collection and Preprocessing Phase

Date	02 June 2024
Team ID	737568
Project Title	AutoForesight : A Predictive Model for Streamlining Car Loan Repayment Planning
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
Kaggle Dataset	Missing Values	Low	Imputation: Replace missing values with a suitable estimate. This could be the mean, median, mode, or a value predicted using machine learning algorithms.
Kaggle Dataset	Duplicate Records	Moderate	Remove Duplicates: Delete duplicate records from the dataset, keeping only one instance of each unique record.

Kaggle Dataset	Categorical data in the dataset	Moderate	Encoding has to be done in the data.