Data Collection and Preprocessing Phase

Date	9 July 2024
Team ID	xxxxxx
Project Title	Human Resource Management Predicting Employee Promotions Using Machine Learning
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.com	Negative Data	Low	Using List comprehension to clear out anomalies ang negative data
Kaggel.com	Categorical Columns	Moderate	Converting education levels to numeric values (1, 2, 3) and encodes the 'department' column as numeric labels using Label Encoder.
Kaggle.com	Imbalanced Data	High	Using SMOTE with 2 nearest neighbors to perform synthetic

			over-sampling on the minority class, creating balanced datasets x_resample and y_resample from the original features x and target y.