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 Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	This machine learning project aims to develop a predictive model for employee promotions using historical HR data. By leveraging various employee attributes such as demographics, job roles, performance metrics, training records, and past promotions, the project seeks to create an objective, data-driven system. The primary objectives are to enhance fairness, reduce biases, streamline the promotion process, improve employee morale, and optimize talent management within the organization.
Data Collection Plan	Kaggle.com

Raw Data Sources Identified	The raw data for this machine learning project comes from the Kaggle dataset "HR Analytics: Employee Promotion Data." This dataset includes employee demographics, education, department, region, gender, recruitment channel, training scores, work experience, previous year ratings, key performance indicators (KPIs), awards won, average training scores, and promotion status. It is designed to help analyze and predict employee promotions. For more details, visit the Kaggle HR Analytics dataset page.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle.com	The dataset includes employee demographics, job information, performance metrics, and promotion status for analyzing and predicting employee promotions.	https://www.kaggle.com/datasets/arashnic/hr-ana	CSV	3.7MB	Public