Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	-
Project Title	Human Resource Management: Predicting Employee Promotions Using Machine Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Report:

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 Report:

Section	Description
Project Overview	This project aims to develop a machine learning model to forecast employee promotions within an organization. By analyzing various factors such as performance metrics, tenure, skills, and feedback, the model will identify high-potential employees who deserve advancement opportunities. This solution enhances workforce management strategies, fosters employee engagement, boosts retention, and promotes organizational growth.
Data Collection Plan	- Extract data from internal HR databases containing employee details, performance metrics, and promotion records. - Prioritize datasets with comprehensive demographic information, including department, education level, and length of service.
Raw Data Sources Identified	The raw data source for this project is a dataset obtained from Kaggle, a popular platform for data science competitions and repositories. The dataset, accessible at Kaggle HR Analytics Dataset, encompasses variables such as employee demographics, education, performance metrics, and more, which are critical for

predicting employee promotions

Raw Data Sources Report

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dateset	The dataset comprises employee details, including demographics, education, performance metrics, and promotion outcomes.	https://www.kaggle.com/datasets/arashnic/hr-ana/data?select=train.csv	CSV	4 MB	Public