**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 15 March 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 project aims to leverage machine learning techniques to predict employee promotions within an organization. The objective is to enhance workforce planning and development by identifying potential candidates for promotion based on various data points.. |
| Data Collection Plan | Data will be collected from internal HR databases, employee performance records, training and development logs, and other relevant sources within the organization. |
| Raw Data Sources Identified | Below are the raw data sources identified for this project, along with relevant details. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Employee Performance Data | Contains records of employee performance reviews and ratings over time. | https://www.kaggle.com/code/janiobachmann/attrition-in-an-organization-why-workers-quit | CSV | 5 MB | Public |
| HR Master Database | Comprehensive HR data including employee demographics, job roles, and tenure. | https://www.kaggle.com/code/klmsathishkumar/hr-analytics-job-change-prediction | CSV | 3 MB | Public |