**Data Collection and Preprocessing Phase**

|  |  |
| --- | --- |
| Date | 23 June 2025 |
| Team ID | xxxxxx |
| Project Title | Credit Card Approval Prediction |
| 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 Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | The machine learning project aims to predict credit card approval based on applicant information. Using a dataset with features such as gender, liability information, loan count, count of EMIs paid off/ existing, no. of employment days, etc. the objective is to build a model that accurately classifies credit card approval eligibility facilitating efficient and informed decision-making. |
| Data Collection Plan | 1. Search for datasets related to credit card applications, financial information and applicant details.  2. Prioritize datasets with diverse demographic information. |
| Raw Data Sources Identified | The raw data sources for this project include datasets obtained from Kaggle , the popular platform for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such as gender, marital status, income, and other related details for machine learning analysis. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Kaggle Dataset 1 | **application\_record.csv** contains appliers personal information, which you could use as features for predicting. | https://www.kaggle.com/code/namphuengauawatcharo/credit-card-approval-prediction/data?select=application\_record.csv | CSV | 54.34 MB | Public |
| Kaggle Dataset 2 | **credit\_record.csv** records users' behaviors of credit card. | https://www.kaggle.com/code/namphuengauawatcharo/credit-card-approval-prediction/data?select=credit\_record.csv | CSV | 15.37 MB | Public |