Data Collection and Preprocessing Phase

Date	10 July 2024
Team ID	740015
Project Title	Credit card approval prediction using ML
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:

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, children, income, and credit history, the objective is to build a model that accurately classifies credit status (approved or denied), facilitating efficient and informed decision-making in the approval process.
Data Collection Plan	• Search for datasets related to credit card approvals, financial information, and applicant details. • Prioritize datasets with diverse demographic information.
Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle & UCI, the popular platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such as gender,

	children, income, and credit card related details for machine learning analysis.

Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset comprises applicant details (gender, children), financial metrics (income, income type), and credit card approval outcomes.	https://www.kaggle.com/namphuengauawatcharo/credit-card-approval-prediction/data	CSV	15 kB	Public
UCI	This data concerns credit card applications; a good mix of attributes	https://archive.ics .uci.edu/dataset/27/credit+approval	CSV	13.6 kB	Public