Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	team-739735
Project Title	Online Payments Fraud Detection
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	The machine learning project aims to detect fraudulent online transactions based on transaction data and user behavior. Using a dataset with features such as transaction amount, time of transaction, location, user demographics, and historical transaction patterns, the objective is to build a model that accurately identifies fraudulent activities. This will facilitate efficient and informed decision-making in fraud prevention, enhancing the security and reliability of online payment systems.
Data Collection Plan	Use platforms like Kaggle and Search for datasets related to Online payments fraud detection.
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 transaction amount, transaction type, origin and destination account details, account balances before and after transactions, and fraud indicators.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle	The dataset comprises factors such as transaction amount, transaction type, origin and destination account details, account balances before and after transactions, and fraud indicators. These variables are critical for identifying patterns and anomalies associated with fraudulent activities in online payments.	https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset	CSV	177MB	Public