

Data Collection and Preprocessing Phase

Date	1 May 2024
Team ID	738286
Project Title	Online Payments Fraud Detection 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	<p>This project aims to develop a system for detecting fraudulent credit card transactions using machine learning algorithms.</p> <p>This project offers a practical approach to building a credit card fraud detection system with potential real-world applications.</p>
Data Collection Plan	<p>In this project we have used PS_20174392719_1491204439457_logs.csv data. This data is downloaded from kaggle.com. Please refer to the link given below to download the dataset.</p>
Raw Data Sources Identified	<p>There are many popular open sources for collecting the data.</p> <p>Eg: kaggle.com, UCI repository, etc.</p>

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	Raw Data of (Type,Amount,Nam eOrig,Old& New BalanceOrg)	Online Payments Fraud Detection Dataset (kaggle.com)	CSV	186 MB	Public