

Data Collection and Preprocessing Phase

Date	14 June 2025
Team ID	SWTID1749662491
Project Title	Online-Payments-Fraud-Detection-using-Machine-Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification 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	Online payment fraud poses a serious threat to customers, disrupting their transaction experience and trust in digital platforms. Users often fall victim to unauthorized transactions or data breaches, especially in the absence of proactive security systems. These incidents lead to dissatisfaction and reluctance to adopt online payment services. To improve user confidence and protect financial data, We aim to detect fraudulent patterns using machine learning. By analyzing transaction behaviors and learning from real-time data, we can deliver a secure, seamless payment experience that aligns with customer expectations and promotes digital trust.

Data Collection Plan	<ul style="list-style-type: none"> Search for datasets related to online payment frauds, financial information, and the type of online transaction. Prioritize datasets with diverse demographic information.
Raw Data Sources Identified	The raw data sources for this project include dataset obtained from Kaggle the popular platform for data science repositories. The provided sample of data represents a subset of the collected information, with the variables such as step, type, amount, nameOrig, oldbalanceOrg, newbalanceOrig, nameDest, oldbalanceDest, newbalanceDest

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset 1	This dataset comprises the details like(step, type, amount, nameOrig, oldbalanceOrg, newbalanceOrig, nameDest, oldbalanceDest, newbalanceDest) and gives the out of whether the online	https://www.kaggle.com/datasets/rupakroy/online-payments-fraud-detection-dataset	CSV	0.4 GB	Public

	transaction is a fraud or not				
--	----------------------------------	--	--	--	--