Technology Stack

Online Payments Fraud Detection using Machine Learning

Technology Stack Overview

Layer	Technology	Version	Purpose
Language	Python	3.10+	Core development language
ML Library	scikit-learn	1.3+	Random Forest training & evaluation
Data Processing	Pandas	2.0+	Data loading, cleaning, EDA
Numerics	NumPy	1.24+	Array operations for inference
Web Framework	Flask	3.0+	Backend web server & routing
Templating	Jinja2	Built-in	Dynamic HTML rendering
Frontend	HTML5 + CSS3	—	User interface & styling
Fonts	Google Fonts	CDN	Caveat + Nunito typefaces
Model Storage	Pickle (.pkl)	Built-in	Serialized Random Forest model
Visualization	Matplotlib / Seaborn	3.7+	EDA charts in Jupyter
Notebook	Jupyter Notebook	7.0+	Model training environment
Dataset	PaySim (Kaggle)	—	Synthetic financial transactions
Version Control	Git / GitHub	—	Source code management

Why These Technologies?

Random Forest Classifier

Chosen over alternatives (Logistic Regression, SVM, Neural Networks) because it: handles class imbalance effectively, provides probability outputs, is interpretable, and achieves high accuracy on tabular data without feature scaling.

Flask

Lightweight Python micro-framework ideal for small ML serving applications. Zero boilerplate compared to Django. Easy to integrate pickle-based models. Jinja2 templating allows dynamic result rendering.

PaySim Dataset

Synthetic dataset generated using agent-based simulation of real mobile money transactions. Contains 6.3 million transactions with ground-truth fraud labels. Widely used benchmark for payment fraud detection research.