

Define Problem Statements

Online Payments Fraud Detection using Machine Learning

IBM SkillsBuild | AI/ML Final Year Project

1. Problem Statement

Online payment fraud represents one of the fastest-growing categories of financial crime globally. Traditional rule-based fraud detection systems suffer from three critical limitations:

1. They cannot adapt to new fraud patterns without manual rule updates
2. They generate high false-positive rates (flagging legitimate transactions as fraud)
3. They operate on static thresholds that fraudsters can learn to exploit

This project addresses the need for an intelligent, data-driven fraud detection system that can learn from historical transaction patterns and generalize to previously unseen fraud attempts.

2. Problem Definition Framework

2.1 Who is Affected?

- End users / customers whose accounts are compromised
- Payment processors (banks, fintech companies) facing financial losses
- E-commerce merchants suffering chargebacks
- Regulatory bodies requiring audit trails and compliance

2.2 What is the Core Problem?

The inability to distinguish fraudulent transactions from legitimate ones in real-time, at scale, with low false-positive rates and high recall on the minority fraud class.

2.3 Why Does It Matter?

- Global payment fraud losses exceeded \$32 billion in 2023
- CASH_OUT and TRANSFER transaction types account for ~100% of fraud in the PaySim dataset
- A 1% improvement in fraud recall can save millions in losses for large payment processors

3. Success Criteria

Metric	Minimum Target	Achieved
Overall Accuracy	$\geq 99\%$	99.96%
Fraud Precision	$\geq 90\%$	100%
Fraud Recall	$\geq 70\%$	75%
F1-Score (Fraud Class)	≥ 0.80	0.85
Web App Response Time	< 2 seconds	< 0.5 seconds

4. Constraints

- Dataset must be publicly available and reproducible
- Solution must run on a standard laptop (no GPU required)
- Web interface must be accessible to non-technical users
- Model must provide a probability score, not just a binary verdict