HackOrbit 2025

Team Adhi

TEAM ADHI

Name : Borigi Jyothiradithya

College : Indian Institute Of Engineering Science and Technology

Department : Computer Science and Technology

Phone Number : 7207180221

E-Mail : adhithyaborigi@gmail.com

Walmart Fraud Detection System

A Machine Learning Based Real-Time-Fraud Detection Web App

Problem Statement :

Challenge:

With the surge in e-commerce transactions, Walmart faces an increasing risk of fraudulent activities such as:

- Fake orders
- Stolen credit cards
- Repeated returns
- Unusual user behavior

Impact:

- Financial losses
- Trust issues with genuine customers
- Operational inefficiencies

PROPOSED SOLUTION :

Objective:

Develop a machine learning-based web application that predicts fraudulent transactions using:

- Transaction behavior
- Customer demographics
- Temporal and velocity features

Solution Highlights:

- Real-time prediction with probabilistic output
- User-friendly Streamlit interface
- Input validation and auto feature transformation

How It Works

User Input:

- Transaction amount, device used, return count, etc.
- Customer location, product category
- Time since last transaction, time of day, etc.

Backend Process:

1. Preprocess input (encoding, dummy vars, reindexing)
2. Load pre-trained RandomForestClassifier
3. Predict fraud probability
4. Return prediction result to user

Explanation of Working Model

◆ **How the Fraud Detection System Works**

Our fraud detection system takes in **user transaction details** and processes them through a trained machine learning model to identify potentially fraudulent activity. Here's how the system operates step by step:

User Input

The system gathers key information related to a transaction, including:

- **Transaction details** such as amount, device used, and return count.
- **Customer attributes** like location and product category.
- **Behavioral patterns** such as time since the last transaction and time of day.

These inputs are crucial for identifying patterns that may indicate fraud.

⚙ **Backend Process**

Once the user input is received, the backend performs the following steps:

1. **Preprocess the input data** – This includes encoding categorical values, generating dummy variables, and reindexing the data to match the model format.
2. **Load the pre-trained model** – Specifically, a **RandomForestClassifier**, which has been trained on historical transaction data.
3. **Predict fraud probability** – The model analyzes the input to determine the likelihood of the transaction being fraudulent.
4. **Return results** – Finally, the prediction is returned to the user in real time, allowing for quick decision-making.

FEATURES

13+

Includes behavior, temporal, and velocity features

Examples:

- `Amount, DeviceID, ReturnCount, CouponUsed`
- `TimeSinceLastTransaction, TransactionsInLastHour`
- `CustomerLocation_x, ProductCategory_x`

Technologies Used :

Programming Language	Python
Web App Framework	Streamlit
Machine Learning	Scikit-learn (Random Forest)
Model Serialization	joblib
Data Handling	pandas
Feature Engineering	Label Encoding, One-Hot Encoding

✕ Drawbacks :

- **1.Data Bias:** Accuracy depends on data quality and balance.
- **2.Outdated Patterns:** Needs regular updates to handle evolving fraud tactics.
- **3.Limited Features:** Doesn’t use advanced signals like IP, payment method, or user history.
- **4.Security Risks:** User data handling needs encryption and authentication.
- **5.Fixed Threshold:** Static 0.5 cutoff may cause misclassifications.

Thank you