🚀 AI-Powered Fraud Detection & Anomaly Analysis Dashboard

# 📌 1. Abstract

This project is an interactive fraud detection system built with Streamlit.

It allows users to upload transaction datasets and automatically detects fraudulent activities or anomalies.

The system supports both supervised learning (with labels) and unsupervised learning (without labels),

making it adaptable to real-world scenarios where fraud labels may or may not be available.

# 📌 2. Objectives

- Automate fraud detection in transaction datasets.

- Provide both supervised (with labels) and unsupervised (without labels) analysis.

- Build an interactive Streamlit dashboard for visualization and decision-making.

- Allow users to preprocess, train, visualize, and download results easily.

# 📌 3. Key Features

✅ Dual-Mode Operation: Supervised (RandomForestClassifier), Unsupervised (IsolationForest)

✅ Smart Preprocessing: Handles numeric, categorical, boolean, missing values automatically

✅ User Control: Upload any CSV, drop irrelevant columns, filter results

✅ Interactive Dashboard: Charts, confusion matrix, classification report, anomaly summary

✅ Result Management: Auto-save results, CSV download for filtered fraud/anomalies

# 📌 4. Workflow / Methodology

1️⃣ Data Upload → CSV file upload & preview

2️⃣ Preprocessing → Drop columns, convert datatypes, build pipeline

3️⃣ Mode Selection → Supervised (labels) or Unsupervised (no labels)

4️⃣ Model Training → RandomForestClassifier / IsolationForest

5️⃣ Visualization → Charts, confusion matrix, anomaly summary

6️⃣ Result Export → Save results, allow CSV download

# 📌 5. Technical Stack

- Frontend: Streamlit

- Data Handling: Pandas, NumPy

- Visualization: Matplotlib, Seaborn

- ML Models: RandomForestClassifier, IsolationForest

- Preprocessing: StandardScaler, OneHotEncoder, ColumnTransformer

# 📌 6. Expected Results

- Detect fraudulent or suspicious transactions

- Provide clear visual insights to decision-makers

- Empower non-technical users to explore fraud datasets

# 📌 7. Future Enhancements

- Add deep learning models (LSTM, Autoencoders)

- Real-time streaming data support (Kafka/Flask API)

- Cloud deployment with authentication

- Explainable AI integration (SHAP values)

# 📌 8. Briefly

“I built an AI-powered fraud detection dashboard that works in both supervised and unsupervised modes.

It preprocesses raw transaction data, trains a model, and provides interactive insights.

For labeled datasets, it uses Random Forest Classifier; for unlabeled datasets, it switches to Isolation Forest.

The dashboard allows users to upload data, drop columns, visualize fraud patterns, and export results.

This makes it a practical tool for real-world fraud analytics in finance and banking.”