# EXPLAINABLE FRAUD ALERT SYSTEM

DANIEL ELLIOT S

DARANI PRASAD M

DAYANIDHI D

# BUSINESS PROBLEM

- Banks use AI to detect suspicious or fraudulent transactions.
- Current systems flag transactions without clear explanations.
- Customers often do not understand why an alert was triggered.
- This causes confusion, loss of trust, and more customer service requests.
- There is a need for a system that explains fraud alerts in simple, user-friendly language.

# CURRENT CHALLENGES

- **Lack of explainability**: Customers don’t understand why a transaction is flagged.
- **Black-box AI models**: High accuracy but no transparent reasoning behind decisions.
- **Customer dissatisfaction**: Confusing alerts lead to frustration and support calls.
- **Delayed response time**: Real-time explanations are difficult with complex models.
- **Trust issues**: Users may lose trust in the bank's fraud detection system due to lack of clarity.

# CURRENT SOLUTIONS

- **XAI Tools (SHAP, LIME):** Explain why transactions are flagged.
- **Behavior Modeling:** Detects unusual spending patterns.
- **LLM-Based Chatbots:** Provide simple, user-friendly explanations.
- **Anomaly Detection:** Uses AI to find and explain suspicious activity.
- **Trust Scoring:** Rates transactions based on risk factors.
- **Visual Dashboards:** Show clear reasons behind fraud alerts.

# OUR SOLUTION

- **Anomaly Detection:** Uses Isolation Forest to identify unusual transactions.
- **Explainable AI (XAI):** SHAP values highlight key features behind fraud flags.
- **Behavioural Modelling:** Captures user transaction patterns for risk analysis.
- **Conversational Interface (Streamlit):** Displays human-readable fraud reasons clearly.
- **Trust Score Indicator:** Each transaction is given a transparency-based risk score.

# Current solution VS Proposed System

| Feature | Current Solutions | Proposed Explainable System |
|---|---|---|
| **Fraud Detection Accuracy** | High (with ML/DL models) | High (uses advanced ML + LLMs) |
| **Explanation to Users** | Minimal or absent | Clear, simple, and user-friendly |
| **Customer Trust** | Low due to opaque decisions | High due to transparent justifications |
| **False Positive Handling** | Moderate control | Improved through behavioral reasoning |
| **User Communication** | Technical alerts or raw flags | Conversational interface with human-like replies |
| **Real-Time Responsiveness** | Limited in complex models | Optimized with lightweight LLM-based explanations |
| **Trust-Based Scoring** | Not typically implemented | Incorporated for personalized fraud risk evaluation |

# BENEFITS

- **Transparency:** Users understand *why* a transaction is flagged using SHAP explanations.
- **Real-Time Fraud Alerts:** Detects suspicious activity instantly using Isolation Forest.
- **User Trust & Confidence:** Trust score and clear explanations build user confidence in the system.
- **Customizable & Scalable:** Easily adaptable to different banking environments and datasets.
- **Interactive Interface:** Streamlit app provides a user-friendly experience for non-technical users.

# ALGORITHMS USED

❖**Isolation Forest**

- Chosen for its efficiency in detecting anomalies without needing labeled data.
- It isolates outliers quickly by randomly splitting features.

❖**SHAP (SHapley Additive exPlanations)**

- Provides clear, human-understandable reasons behind each fraud alert.
- Enhances model transparency and builds user trust.

❖**MinMaxScaler**

- Ensures all features are on the same scale, improving model accuracy.
- Important for distance-based models like Isolation Forest.

# RESULTS & IMPACTS

**Results:**

- Accurate Anomaly Detection
- Clear Explanations
- Real-Time Processing
- Interactive Dashboard

**Business Impact:**

- Enhanced Fraud Prevention
- Increased Customer Trust
- Lower Operational Costs
- Scalable Risk Solution

# THANK YOU