

# **EXPLAINABLE FRAUD ALERT SYSTEM**

**Behavior-Based Detection + Gemini AI  
Explanations**

Presented by: Rohith K ( 262)

Sabesh V(263)

Sajjan Kumar (264)

Sakthi priyan(265)



# ABSTRACT

- Financial fraud is increasing and needs smarter detection.
- Traditional models lack explanation, leading to user confusion.
- Our system uses ML for behavior modeling and Gemini AI for simple explanations.
- Goal: Transparent fraud detection with user trust.



# PROBLEM STATEMENT

01.

Bank flags suspicious transactions, but users don't know why

02.

Users need clear, understandable fraud reasons.

03.

Black-box models lack of trust, false complaints.

04.

Solution: Explainable AI with real-time feedback.

# OBJECTIVES

- Model normal transaction behavior per user.
- Detect anomalies with Isolation Forest.
- Use Gemini AI to generate natural language explanations.
- Assign trust scores for flagged transactions.
- Display results in a simple web UI.



# TECHNOLOGIES STACK



- Python 3.12 – Programming
- Scikit-learn – ML model (Isolation Forest)
- Gemini Pro API – LLM explanation
- Streamlit – Interactive web app
- Pandas/Matplotlib – Data handling and plotting
- FPDF – PDF report generation

# HOW IT WORKS

01

Upload transaction  
history (CSV)

02

ML model detects  
unusual behavior

03

Gemini generates  
human-friendly  
reason

04

Trust score is assigned,  
Output is shown in UI  
(highlighted & ranked)

# SYSTEM ARCHITECTURE FLOW

## Input Layer

- **Data Source:** Transaction history (CSV format) containing details like transaction amount, date, time, merchant, and user behavior.
- **Process:** User uploads CSV through the Streamlit Interface.

## Anomaly Detection (Isolation Forest)

- **Algorithm:** Isolation Forest (a machine learning model) processes the transaction data.
- **Function:** Identifies transactions that deviate from a user's normal spending behavior.
- **Result:** Suspicious transactions are flagged for further processing.

## Gemini NLP Explanation

- **Input:** The flagged transactions and their corresponding behaviors.
- **Function:** The Gemini Pro API generates natural language explanations for the flagged transactions, providing users with understandable reasons for the alerts

## Trust Score System

- **Input:** Each flagged transaction and its context (e.g., amount, frequency).
- **Process:** A trust score is calculated based on the risk level of the transaction.
- **Output:** A numerical score that indicates the likelihood of the transaction being fraudulent (e.g., 90% trust = high risk).

## Streamlit Interface (UI)

- **User Interaction:** The user uploads the CSV, and results (flagged transactions, explanations, and trust scores) are displayed.
- **Feature:** Results are visualized in a simple, interactive web interface with highlighted flagged transactions and the corresponding explanations.
- **Action:** Users can view detailed reasons for flagged transactions and prioritize them based on trust scores.

# KEY FEATURES

- Accurate anomaly detection
- Human-readable NLP explanations
- Risk-level trust scores
- Clean, responsive UI
- Exportable PDF reports
- Free Gemini API usage = cost-effective

# RESULTS & BENEFITS

- Outliers detected with high precision
- Explanations help reduce user complaints
- Transparent fraud alerts      higher customer trust
- Interface enables quick review by analysts

# CONCLUSION & FUTURE WORK

## Conclusion:

- Successfully combined ML & Gemini AI for fraud detection with explanations.
- Reduces confusion, false positives, and boosts trust.

## Future Plans:

- Real-time deployment with bank APIs
- Add support for regional languages
- Visual dashboard with live monitoring

**95% Model Accuracy**

**Thank you !**