deriv

AI-Powered Client Fraud Detection & Investigation

The Challenge

How might we build an AI system that detects, investigates, and explains suspicious client activity - from identity fraud to money laundering - while reducing false positives and empowering investigators to work faster and smarter?

The Problem

Fraud teams face an impossible task: review thousands of alerts daily, investigate suspicious accounts across multiple systems, distinguish real fraud from false positives, and document everything for regulators - all while fraudsters evolve faster than rule-based systems can adapt.

> *"I review 200 alerts per day. Maybe 10 are real fraud. But I can't skip any because the one I skip might be the critical one."*
>
> *"The system flags an account as 'high risk' with a score of 87. But it doesn't tell me WHY. I spend an hour gathering data from five different systems just to understand what triggered it."*
>
> *"By the time I investigate one case, ten more pile up. And half of them are false positives that look identical to real fraud."*
>
> *"We discovered a new fraud pattern last month. Looking back, it had been happening for six months - we just didn't see it until losses got big enough to notice."*
>
> *"A student declares $500 monthly income but deposits $8,000 in two weeks. An account uses AI-generated faces for verification. Someone accesses from a sanctioned country via VPN. These all require investigation, but I don't have enough hours in the day."*

The core issue: investigators spend 80% of their time on data gathering and noise, 20% on actual fraud detection and decision-making. Meanwhile, sophisticated fraud hides in plain sight.

Why This Matters Now

Financial platforms face increasingly sophisticated client-side fraud:

- Identity fraud: Synthetic IDs, deepfakes, stolen documents, AI-generated faces
- Money laundering: Fake trades, rapid deposit-withdrawal cycles, transactions inconsistent with customer profiles
- Account takeover: Credential stuffing, phishing, social engineering
- Behavioural anomalies: Deposits far exceeding declared income, unusual geographic access patterns
- Sanctions evasion: VPN usage to mask location, accessing from prohibited jurisdictions

Current challenges:

- Alert overload: Rule-based systems generate 95%+ false positives

- Manual investigation: Data gathering takes hours per case
- Reactive detection: Fraud patterns discovered months after they start
- Inconsistent decisions: Different investigators interpret the same signals differently
- Documentation burden: Writing investigation reports is time-consuming
- Evolving threats: Fraudsters adapt faster than static rules can be updated

Fraud teams need AI that detects, explains, prioritises, investigates, and learns - continuously.

The Opportunity

Build an integrated AI fraud detection and investigation system with four core capabilities:

1. Intelligent Detection & Pattern Discovery:
    - Multi-signal detection: Monitor identity verification, transaction behaviour, geographic access, and account activity
    - Behavioural anomaly detection: Flag deposits inconsistent with declared income, unusual transaction patterns, geographic impossibilities
    - Identity fraud detection: Detect forged documents, synthetic IDs, AI-generated faces, deepfakes, biometric mismatches
    - Sanctions & geographic risk: Identify VPN usage, access from prohibited jurisdictions, and location anomalies
    - Emerging pattern discovery: Use unsupervised learning to discover new fraud typologies before they become widespread
    - Network analysis: Connect related accounts through shared IPs, devices, timing patterns, or coordinated behaviour

2. Explainable Alerts & Prioritisation:
    - Clear explanations: Generate human-readable explanations - "Flagged because: deposits 1400% above declared income + 73% of logins from high-risk jurisdiction via VPN"
    - Confidence scoring: Risk levels with reasoning - why is this 87 vs. 65?
    - Intelligent prioritisation: Learn from historical outcomes to rank alerts by likelihood of being real fraud
    - False positive prediction: Identify and auto-resolve obvious false positives with audit trail
    - Contextual comparison: "This pattern matches 12 confirmed money laundering cases from last quarter"

3. AI-Assisted Investigation:
    - Automated data gathering: Pull relevant information from multiple systems into a coherent case view in seconds
    - Timeline reconstruction: Build chronological sequences of suspicious activity automatically
    - Evidence synthesis: Summarise transaction history, login patterns, document verification results, and behavioural signals
    - Investigation suggestions: Recommend next steps - "Check if this IP is shared with other accounts"
    - Cross-case connections: "This device was used by 7 other flagged accounts"

4. Documentation & Learning:

- Automated report generation: Create investigation summaries and audit-ready compliance documentation
- Regulatory explanations: Generate findings suitable for different audiences (investigators, compliance, regulators)
- Continuous learning: Improve detection accuracy based on investigator feedback and confirmed outcomes
- Knowledge capture: Learn from every investigation to improve future detection

The system should turn a 2-hour investigation into a 10-minute review while catching fraud that manual processes miss.

Constraints

Constraint	Rationale
Must demo live	Show working software, not concept slides.
AI must add value	This is an AI hackathon. GenAI must be core to your solution.
Human in the loop	AI assists and recommends; investigators make final decisions.
Explainable decisions	Every alert and recommendation must have clear reasoning.
No missed fraud	False positive reduction cannot sacrifice fraud detection.

Questions Worth Considering

- What's the most time-consuming part of fraud investigation that AI could accelerate?
- How do you balance comprehensive detection with manageable alert volumes?
- What makes an explanation actionable vs. just informative?
- How do you distinguish emerging fraud patterns from random anomalies or legitimate business changes?
- When should AI suggest next steps vs. wait to be asked?
- How do you handle cases where multiple fraud signals conflict?
- Can you connect dots across identity, transactions, and behaviour that investigators wouldn't see?
- How do you present complex case information clearly without overwhelming the investigator?

What Would Blow Our Minds

- Investigation copilot: Complete case summary with evidence, timeline, and recommendations in 30 seconds
- Pattern discovery: Finding fraud techniques weeks before they become obvious - "Detected emerging behaviour cluster: 47 accounts with similar transaction patterns"
- Network revelation: "This account is part of a 23-account fraud ring sharing 3 devices and coordinating transaction timing"
- Predictive explanations: "This alert is 89% likely to be real fraud based on 5 strong signals matching confirmed cases"
- Learning from outcomes: A System that improves daily based on investigator decisions and discovered patterns
- Cross-domain intelligence: "Identity documents passed automated checks, but transaction behaviour + VPN usage + device fingerprint all match known fraud network"

- Audit-ready documentation: Investigation reports generated automatically that meet regulatory standards
- 5x productivity: Investigators handling 5x more cases with higher accuracy