AI Security Council - Hackathon Project Report

Team Name : Deadlock

Members : Khaled Akra , Sirine Khalil , Mariam Khalil

42Beirut x Teknologiia Hackathon
Date November 16, 2025

Executive Summary

I built an AI Security Council system that uses three different AI models (DeepSeek and Gemini) working together to analyze cybersecurity threats. The goal was to reduce false positives in SOC alerts by having multiple AI agents verify each threat before taking action.

Tech Stack :
- Backend : N8N workflow automation (Docker)
- AI Models : DeepSeek (Agent 1), Gemini 2.5 Flash(Agent 2), Gemini 2.5 Flash (Agent 3)
- APIs : IPInfo, AlienVault OTX, IP-API
- Frontend : HTML/CSS/JavaScript dashboard
- Infrastructure : Docker Compose with 3 containers

Final Result :
- 3 AI agents analyzing 10 threats in parallel
- True Positive / False Positive classification
- Real-time dashboard with auto-refresh
- Geographic threat distribution
- Strategic recommendations for SOC teams

The Procedure: How It Works

1. System Architecture

The system follows a three-tier analysis pipeline:

Agent 1 - TRIAGE (DeepSeek)
- Reads threats from CSV file (IP, port, attack type, risk score)
- Performs rapid risk assessment
- Outputs: URGENT / INVESTIGATE / ROUTINE

Agent 2 - HUNTER (Gemini 2.5 Flash)

- Takes Agent 1's output
- Queries 3 threat intelligence APIs in parallel:
- IPInfo: Geolocation and organization data
- AlienVault OTX: Threat pulse information
- IP-API: Additional geolocation verification
- Performs deep analysis with TP/FP classification
- Outputs: THREAT / SUSPICIOUS / BENIGN + TRUE_POSITIVE / FALSE_POSITIVE

Agent 3 - ORACLE (Gemini)

- Receives combined data from Agents 1 & 2
- Performs strategic pattern analysis
- Adds geopolitical context assessment
- Outputs final recommendation: BLOCK / ESCALATE_TO_SOC / MONITOR / ALLOW

Aggregation & Dashboard

- Combines all 10 threat assessments
- Calculates statistics (TP count, FP count, accuracy rate)
- Builds geographic distribution map
- Generates executive recommendations
- Saves to JSON and displays on dashboard

2. Data Flow

```
CSV Input (10 threats)
  □
Agent 1 (DeepSeek) □ Risk scoring
  □
Agent 2 (Gemini 2.5 Flash) □ APIs (IPInfo, AlienVault, IP-API) □ TP/FP classification
  □
Agent 3 (Gemini 2.5 Flash) □ Strategic analysis + Geopolitical context
  □
Aggregation Node □ Statistics calculation
  □
Dashboard (Auto-refresh every 5s)
```

3. Key Features Implemented

True Positive / False Positive Detection

- Each threat gets classified as TP or FP
- Dashboard shows: 2 True Positives, 8 False Positives, 80% FP Rate, 20% Accuracy
- Reduces SOC alert fatigue by filtering out benign traffic

Geographic Analysis

- Top 5 countries by threat count

- Example: United States: 3, China: 2, France: 1, Switzerland: 1, India: 1

Priority Threats Section
- Shows only THREAT verdicts or ESCALATE_TO_SOC recommendations
- Full geopolitical reasoning (no truncation)

Strategic Recommendations
- Immediate Actions: "Block 2 high-confidence threats immediately"
- Short-term: "Monitor 9 suspicious IPs for 24-48 hours"
- Long-term: Strategy for reducing false positive rate

Final Statistics

Total Development Time : 21 hours
Lines of Code Written :
- N8N Workflows: ~800 lines (JavaScript in Code nodes)
- Dashboard: ~900 lines (HTML/CSS/JavaScript)
- Backend API: ~150 lines (Python)

API Calls Per Workflow Run :
- 10 DeepSeek calls (Agent 1)
- 30 HTTP requests (Agent 2: 10 IPInfo + 10 AlienVault + 10 IP-API)
- 10 ChatGPT/Gemini calls (Agent 2)
- 10 Gemini calls (Agent 3)
- Total: 70 API calls per complete analysis

What I Learned

1. N8N is powerful for rapid prototyping - Built complex AI orchestration without writing backend code
2. API format compatibility is critical - Switching AI providers requires understanding their JSON structures deeply
3. Real-time dashboards need polling - Auto-refresh via setInterval is simple and effective
4. Multi-AI consensus reduces false positives - Having 3 agents vote significantly improved accuracy

This project demonstrates how multiple AI models can collaborate to solve real-world cybersecurity challenges, specifically reducing SOC analyst burnout from false positive alerts.