Gen AI-Based Data Profiling Solution

Introduction

Regulatory reporting in the banking sector involves compiling vast amounts of data to meet compliance requirements.

Data profiling is a critical part of this process, ensuring that reported data aligns with regulatory requirements.

Traditionally, this process is manual and time-consuming. This solution automates data profiling using Generative AI (LLMs) and unsupervised machine learning.

Demo

- Input: CSV file with transaction data.

- Processing: Profiling rules are applied, risk scores are calculated, and violations are detected.

- Output: Profiling results, risk scores, and suggested remediations.

Inspiration

Manual data profiling is labor-intensive and prone to human error. Automating this process using AI reduces errors, increases efficiency, and ensures compliance with complex regulatory requirements.

What It Does

- Generates profiling rules based on transaction data and banking regulations.

- Calculates adaptive risk scores using unsupervised machine learning techniques.

- Detects potential violations and suggests remediation actions using a chatbot.

How We Built It

- Backend: Java Spring Boot (Gradle) with RESTful APIs.

- Frontend: React.js with state management.

- Profiling Engine: LLM-based rule generation and risk scoring.

- Chatbot: Suggests remediation actions based on profiling results.

Challenges We Faced

- Handling large transaction data efficiently.

- Defining accurate profiling rules.

- Generating meaningful remediation suggestions using LLMs.

Sample Test Data (Extended)

TransactionID	Amount	Currency	Country	Status	AccountType	TransactionType	Date
T001	15000	USD	US	Completed	Savings	Transfer	2025-03-10
T002	-500	EUR	DE	Failed	Checking	Withdrawal	2025-03-12
T003	3000	GBP	UK	Pending	Savings	Deposit	2025-03-14
T004	12000	USD	US	Completed	Checking	Transfer	2025-03-16
T005	-100	EUR	FR	Failed	Savings	Withdrawal	2025-03-18
T006	20000	USD	US	Completed	Savings	Transfer	2025-03-20
T007	0	JPY	JP	Failed	Checking	Deposit	2025-03-22
T008	2500	EUR	IT	Pending	Savings	Transfer	2025-03-24
T009	-1200	GBP	UK	Failed	Checking	Withdrawal	2025-03-26
T010	8000	USD	CA	Completed	Savings	Deposit	2025-03-28

Solution Output

TransactionID	Risk Score	Violation	Suggested Remediation

T001	1.0	Large transaction	Review and verify
T002	1.0	Negative amount	Check for errors or fraud
T003	0.1	No violation	No action needed
T004	1.0	Large transaction	Review and verify
T005	1.0	Negative amount	Check for errors or fraud
T006	1.0	Large transaction	Review and verify
T007	0.9	Zero amount transaction	Verify with customer
T008	0.2	No violation	No action needed
T009	1.0	Negative amount	Check for errors or fraud
T010	0.5	Unusual transaction for customer	Verify details