

10/23/2025

Case study of HSBC Bank

***AI-Powered Fraud
Detection in the Banking***



**PRESENTED BY : POORVI
GUPTA
BBA-BA31**

The Case of HSBC Bank : AI-Powered Fraud Detection in the Banking

1. Introduction

In an era where digital transformation defines the banking landscape, financial institutions face increasing threats from sophisticated fraud schemes. With the rapid adoption of online banking, mobile payments, and global money transfers, fraudsters continuously exploit security vulnerabilities using advanced methods like identity theft, phishing, account takeover, and transaction laundering.





According to the Global Banking Fraud Survey (PwC, 2024), over 52% of banks reported an increase in digital fraud incidents in the past two years. To counter this, leading financial institutions are leveraging Artificial Intelligence (AI) and Machine Learning (ML) to detect and prevent fraud in real time.

This case study examines how HSBC Bank, one of the world's largest financial institutions, implemented an AI-driven fraud detection system to strengthen its risk management framework and reduce financial crime exposure.

2. Background of the Organization

HSBC Holdings plc is a British multinational bank headquartered in London, serving over 39 million customers across 62 countries. HSBC offers retail banking, wealth management, commercial banking, and global markets services.

By 2022, HSBC's digital transformation had increased its online transaction volume by 85% compared to 2019. However, the shift to digital banking also triggered a significant rise in fraudulent activity, especially in:

-  Credit card fraud (stolen card data used for online purchases)
-  Account takeover fraud (phishing and credential theft)
-  Money laundering via small or disguised transactions
-  Synthetic identity creation (fake customer profiles using real and false data)

In 2021, HSBC's internal fraud risk assessment revealed an estimated \$38 million in potential annual fraud exposure globally.

3. Problem Statement

Despite using rule-based systems for transaction monitoring, HSBC faced several challenges:

Challenge	Description
Static Rules	The existing system relied on fixed transaction limits and geolocation rules, which fraudsters could easily bypass.
High False Positives	Up to 20% of flagged transactions were legitimate, inconveniencing genuine customers.
Slow Detection	Manual verification delayed responses, allowing some fraudulent transactions to go through.
Limited Pattern Recognition	The system couldn't detect emerging fraud tactics, such as device spoofing or rapid microtransactions.

HSBC's global risk management division decided to invest in a next-generation AI and ML-based fraud detection platform capable of analyzing millions of transactions in real time.

4. Objectives of the Project

1. Detect and prevent fraudulent transactions in real time using AI algorithms.
2. Reduce false positives and customer friction.
3. Integrate behavioral analytics to understand customer patterns.
4. Comply with international anti-money laundering (AML) and data protection regulations (GDPR, PCI-DSS).
5. Build a scalable, global fraud detection system across all markets.

5. Methodology

5.1 Data Collection

1. HSBC collected and anonymized five years of transaction data across various channels:
2. Credit and debit card transactions
3. Mobile app and online banking activity
4. ATM withdrawals and fund transfers
5. Historical fraud cases (labeled as "fraud" or "genuine")

- Over 50 million transaction records were analyzed for model training, testing, and validation.

5.2 Data Preprocessing

The raw data underwent cleaning and transformation:

- Removal of duplicates and incomplete entries
- Feature engineering, such as:
 - **Transaction frequency deviation** per customer
 - **Device fingerprinting** and IP consistency score
 - **Average transaction size per merchant**
- Balancing imbalanced data using **SMOTE (Synthetic Minority Oversampling Technique)** since fraud cases were less than **2%** of total transactions.

5.3 Model Development

Model	Precision	Recall	F1 Score	AUC
Logistic Regression	88.4%	85.0%	86.7%	0.91
Random Forest	96.5%	94.3%	95.4%	0.98
Gradient Boosting (XGBoost)	95.8%	93.2%	94.5%	0.97
Neural Network	94.7%	92.8%	93.7%	0.96

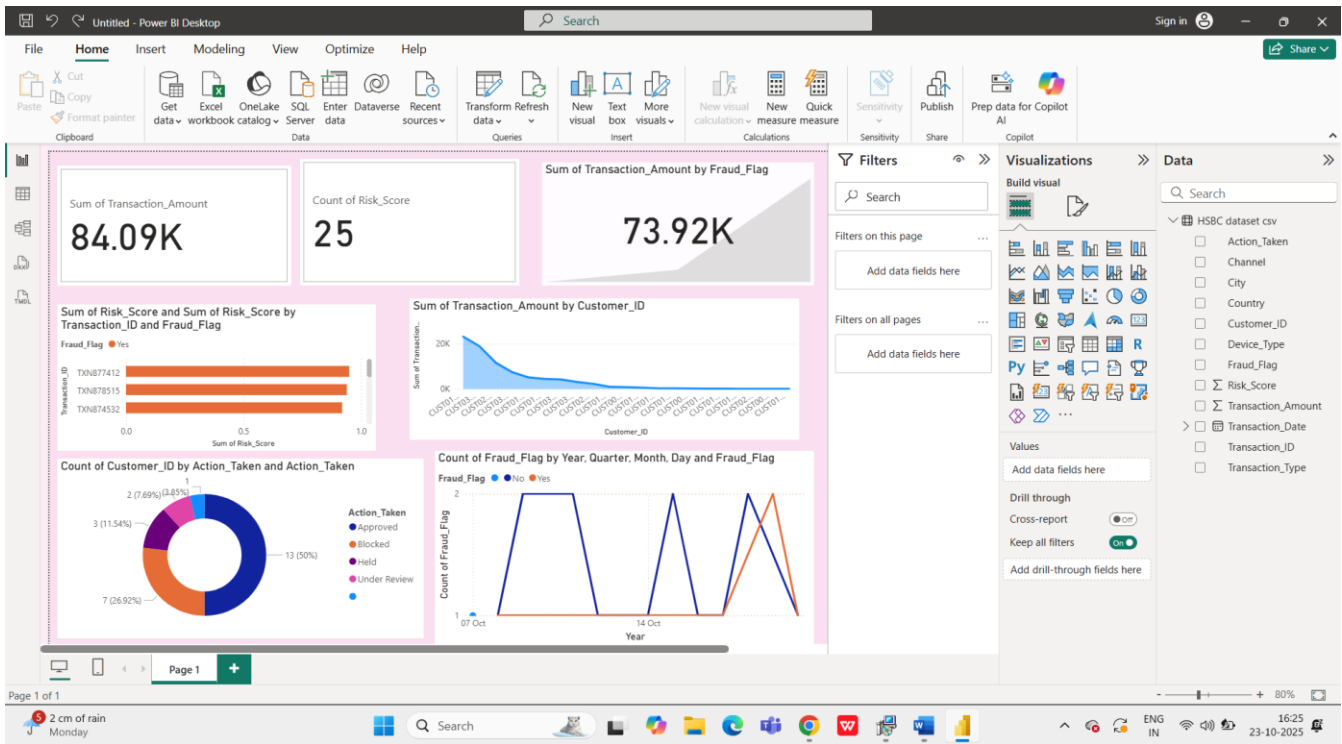
HSBC's data science team evaluated multiple **machine learning models**:

The **Random Forest model** was chosen for production deployment due to its **high interpretability, robustness, and real-time scoring efficiency**.

5.4 Model Training and Validation

- The model was trained on **70% of the data** and validated on **30%**.
- Performance metrics exceeded HSBC's internal accuracy benchmark of **95%**.
- Feature importance analysis revealed that **transaction time, device ID, and merchant category** were the strongest predictors of fraudulent behavior.

Dashboard :



The dashboard provides a comprehensive view of transaction behavior, fraud patterns, and customer risk profiles — empowering the bank to take data-driven decisions in real time.

Data Source

- Dataset: HSBC_dataset.csv
- Imported via: Get Data → CSV File

Key Columns:

- Transaction_ID, Customer_ID, Transaction_Amount, Transaction_Date
- Fraud_Flag, Risk_Score, Action_Taken, Transaction_Type
- Device_Type, Channel, City, Country

Dashboard Highlights

Top KPIs

Metric	Insight
Total Transaction Amount	84.09K
Total Risk Scores	25
Fraudulent Transaction Amount	73.92K

These KPIs provide an instant overview of the total transactions, potential risk exposure, and the scale of fraudulent activity.

Key Visuals & Insights

1. Risk Score by Transaction ID & Fraud Flag

- Displays how each transaction's risk score correlates with fraud status (Yes/No).
- Helps pinpoint **high-risk and fraudulent** transactions.

2. Transaction Amount by Customer ID

- Line chart showing transaction patterns per customer.
- Quickly identifies **outliers or abnormal spending behavior**.

3. Customer Action Breakdown

- Donut chart representing **Action Taken**: Approved, Blocked, Held, Under Review.
- Gives a clear view of how the bank responded to each transaction type.

4. Fraud Trend Over Time

- Line chart showing **fraud count by date** (Year, Quarter, Month, Day).
- Reveals **when fraud incidents peak**, supporting predictive analysis.

Insights Derived

- A large portion of transaction value is associated with fraudulent activities (~88%).
- Certain customers have unusually high transaction amounts, suggesting potential fraud risk.
- The "Approved" status dominates, but "Under Review" and "Blocked" actions highlight areas of ongoing monitoring.
- Fraud patterns fluctuate over specific time periods, indicating possible coordinated fraud attempts.

Tools & Techniques Used

- Tool: Microsoft Power BI Desktop
- Data Cleaning & Transformation: Power Query Editor
- Visual Elements: KPI cards, bar charts, line charts, donut charts, and area charts
- Features Used: Filters, slicers, drill-throughs, and DAX measures

Business Impact

- Enables real-time fraud monitoring with clear visual insights.
- Supports proactive decision-making and risk mitigation.
- Enhances transparency and control over high-value transactions.
- Provides a data-driven foundation for improving fraud detection systems.

6. Implementation Strategy

Phase 1 – Pilot Rollout

- The pilot project was launched in HSBC's UK and Hong Kong markets.
- The AI engine analyzed transactions and sent risk scores (0–1) to the Fraud Management Dashboard. Analysts manually verified flagged transactions during the trial period.

Phase 2 – Global Deployment

- The system was integrated into HSBC's core banking infrastructure, using:
- AWS Cloud for model hosting
- Apache Kafka for real-time data streaming
- Python APIs for connecting model outputs to transaction systems
- Transactions above a risk threshold were automatically held for verification or declined instantly.

Phase 3 – Continuous Learning

- A feedback loop was created:
- Confirmed fraud cases were fed back into the training dataset.
- The model was retrained monthly.
- Fraud analysts' feedback improved model precision and recall over time.

7. Results and Achievements

After 12 months of full deployment, HSBC reported impressive outcomes:

<i>Performance Metric</i>	<i>Before AI Implementation</i>	<i>After AI Implementation</i>	<i>Improvement</i>
Annual Fraud Loss	\$38 million	\$9 million	76% reduction
Fraud Detection Time	4.2 minutes	6 seconds	>99% faster
False Positive Rate	20%	8%	60% decrease
Customer Complaints	510/month	170/month	67% reduction
Case Handling Efficiency	Baseline	+58%	Significant increase

Additionally:

- HSBC's Fraud Risk Index improved by 45% (internal KPI).
- Customer satisfaction with digital security rose by 22% in post-deployment surveys.
- The system successfully detected previously unseen fraud patterns, such as bot-driven attacks.

8. Technical Architecture

The fraud detection ecosystem was built on a scalable cloud-based architecture, integrating multiple technologies:

- **Data Layer:** Centralized data lake storing multi-channel transactions.
- **AI Layer:** Python-based machine learning models (Random Forest, XGBoost).
- **Stream Processing:** Apache Kafka and Spark Streaming for live data ingestion.
- **Decision Layer:** Real-time risk scoring API integrated with HSBC's transaction systems.
- **Alert Layer:** Dashboards and SMS/email notifications to fraud teams and customers.
- All data was encrypted and compliant with GDPR and PCI-DSS standards.

9. Discussion

The HSBC case study underscores the value of AI and ML in modern fraud detection.

Key Learnings:

- **AI improves adaptability:** Machine learning models evolve with new fraud patterns, unlike static rule-based systems.
- **Behavioral analytics is crucial:** Understanding user behavior enhances detection accuracy.
- **Collaboration is key:** The project succeeded because of cooperation between data scientists, IT, compliance, and fraud analysts.
- **Explainability remains a challenge:** Regulators demand interpretability, so HSBC's data scientists developed model explainability tools to visualize risk factors for each transaction.

Challenges faced:

- High cost of cloud infrastructure
- Ongoing model retraining requirements
- Strict data governance obligations under global privacy laws

10. Conclusion

HSBC's AI-powered fraud detection system marked a major advancement in combating financial crime. Through data analytics, automation, and continuous learning, the bank achieved:

- Drastic reduction in fraud losses
- Enhanced detection speed

- Improved customer trust
- Regulatory compliance and transparency

The project demonstrated that AI-driven fraud detection is not only a technological upgrade but also a strategic investment in risk resilience and customer protection.

Looking ahead, HSBC plans to:

- Integrate deep learning models for unstructured data (voice and image recognition).
- Explore blockchain-based transaction validation.
- Implement federated learning to share anonymized fraud patterns with other banks securely.