Credit Card Fraud Detection

A Business Analytics Project

Group 25

Tran Quang Trong 20235565

Pham Quoc Thai 20235558

Nguyen Khac Viet Anh 20235471

Le Thanh Vinh 20200668

December 20, 2025

1 Introduction

1.1 Business Context

In the current era of digital finance, credit card transactions have become the primary method of global commerce. While this shift offers convenience, it also exposes financial institutions to increasingly sophisticated fraudulent schemes. Credit card fraud does not only result in direct monetary theft but also undermines the core of the banking business: customer trust. To remain competitive and secure, banks must transition from reactive, manual fraud detection processes to a proactive risk management framework that can categorize and respond to transaction threats in real-time, safeguarding both institutional assets and the integrity of the payment ecosystem.

1.2 Problem Statement

The institution's current fraud prevention infrastructure is struggling with three strategic hurdles:

- **Extreme Data Imbalance:** Fraudulent activities account for a tiny fraction (approximately 0.17%) of total transactions. This "needle-in-a-haystack" scenario causes traditional systems to overlook subtle fraud patterns, leading to significant annual financial leakage.

- **Customer Friction:** Static and rigid rule-based systems often trigger "False Positives," mistakenly declining legitimate transactions from high-value customers. This results in a poor user experience, loss of transaction revenue, and long-term damage to customer loyalty.
- **Operational Inefficiency:** The sheer volume and velocity of modern transactions have overwhelmed manual review teams. Without an automated risk-prioritization mechanism, investigators waste resources on low-probability cases while missing high-stakes fraudulent strikes.

1.3 Stakeholder Objectives

The success of this project is defined by its ability to fulfill the specific requirements of key business stakeholders:

1. **Risk Management Department:** To maximize the detection rate of fraudulent transactions, ensuring that high-risk threats are identified and neutralized before settlement occurs.
2. **Customer Experience (CX) Team:** To minimize the false alarm rate, ensuring that "good" customers enjoy a seamless and frictionless payment journey without unnecessary interruptions.
3. **Operations and Finance:** To automate the risk classification process, thereby optimizing human capital and reducing the operational costs associated with manual fraud investigation.
4. **IT and Compliance:** To establish a scalable and secure analytical framework that adheres to data privacy regulations while remaining ready for seamless integration into the existing banking infrastructure.

1.4 Business Objectives

The project aims to transform raw data into a strategic risk management asset by achieving the following objectives:

1. **Financial & Operational Optimization:** Minimize the Total Cost of Fraud by balancing direct financial losses (leakage) against the operational costs of manual investigations. The goal is to maximize fraud detection (Recall) while strictly limiting False Positives to ensure high approval rates for legitimate customers, thereby satisfying both Risk and CX stakeholders.

2. **Dynamic Behavioral Intelligence:** Identify temporal and behavioral signatures of fraud. By differentiating "normal" vs. "anomaly" patterns, the system provides actionable insights into evolving criminal tactics rather than relying on static, outdated rules.

3. **Scalable Deployment & Decision Support:** Build a lightweight, high-performance detection engine designed for easy integration into production APIs. The solution is packaged to allow operational teams to prioritize high-risk alerts effectively, bridging the gap between complex modeling and real-time business decisions.

2 Data Description and Preparation

2.1 Dataset Overview

The analysis utilizes a dataset of 284,807 European credit card transactions. From a business risk perspective, the data exhibits an extreme class imbalance, with only 492 transactions (0.172%) identified as fraudulent. While the 28 PCA-transformed features ($V1 - V28$) remain anonymized for confidentiality, the preparation process focuses on the *Time* and *Amount* fields to derive actionable behavioral patterns.

2.2 Temporal Split Strategy

To simulate a real-world banking production environment, we implemented a temporal split instead of random sampling. The first 24 hours of data (containing 269 frauds) were assigned to the training set, while the subsequent 24 hours (containing 223 frauds) formed the test set. This ensures the model is evaluated on its ability to predict "future" fraud based on "past" historical patterns, mimicking live operational conditions.

2.3 Data Quality and Robust Scaling

The dataset is of high quality with no missing values. However, to manage the high skewness of transaction amounts and protect the model from being distorted by high-value outliers (e.g., luxury purchases), we applied the *RobustScaler*. By using median and interquartile range (IQR), this method ensures that the scoring logic remains stable across diverse spending profiles. The scaling parameters are persisted to maintain consistency between the analytical phase and the final deployment interface.

3 Feature Engineering and Exploratory Analysis

3.1 Behavioral Patterns in Transaction Value

Initial analysis reveals that transaction amounts are extremely right-skewed, with the vast majority of activities involving small denominations. From a risk management perspective, two key insights emerge:

- **Camouflage Tactics:** Fraudsters do not exclusively target high-value transactions. Instead, a significant portion of fraud occurs at low-to-moderate amounts to blend into "normal" daily spending and avoid triggering limit-based alerts.
- **Inadequacy of Simple Rules:** Relying solely on transaction size as a fraud indicator is insufficient. A robust detection strategy must look beyond absolute values and focus on the relationship between amount and other behavioral dimensions.

3.2 Risk-Time Correlation

The raw temporal data was transformed into an *hour-of-day* format to better capture the cyclical nature of consumer behavior and criminal activity:

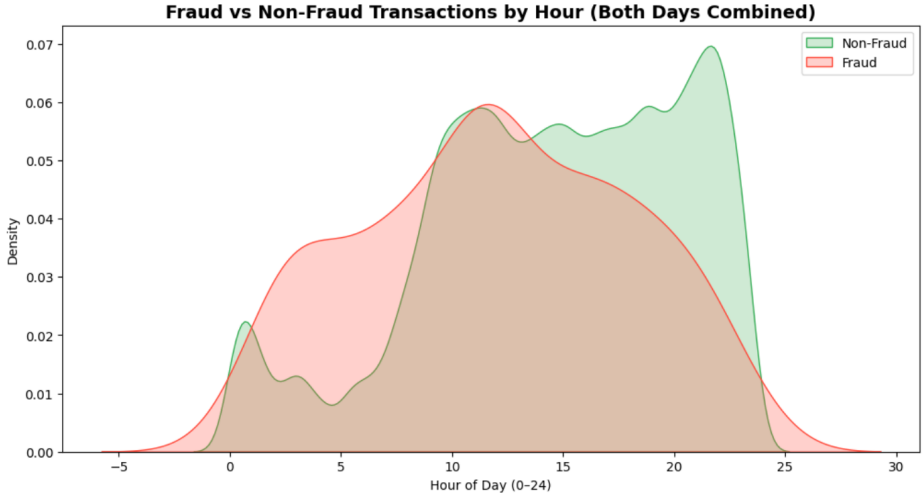

Figure 1:

- **Off-Peak Vulnerability:** Fraudulent density increases significantly between 1 AM and 5 AM—a window where legitimate activity is at its lowest. This represents a period of "High Risk" due to potentially slower manual review response times.

- **Peak-Hour Camouflage:** During business hours (8 AM - 8 PM), fraud activity mirrors legitimate traffic. This suggests that criminals strategically synchronize their strikes with high-volume periods to hide within the "noise" of heavy transaction traffic.

3.3 Strategic Implications

These findings transition the project from simple data observation to **Targeted Monitoring**. By identifying these high-risk windows and value-ranges, the bank can optimize its resource allocation: increasing automated scrutiny during off-peak hours and deploying more sophisticated behavioral filters during peak periods to maintain a frictionless experience for legitimate users.

3.4 Correlation Insights and Signal Selection

A correlation analysis of the PCA-transformed features was conducted to identify the most potent indicators of fraudulent activity. The key business insights include:

- **Feature Independence:** The components $V1 - V28$ exhibit near-zero correlation with each other since these features are used by PCA.
- **High-Impact Risk Indicators:** A specific subset of features demonstrates a significantly stronger correlation with the fraudulent class, including:

$$V3,\ V7,\ V10,\ V12,\ V14,\ V16,\ V17$$

- **Targeted Anomaly Detection:** These "high-impact" components act as primary signals for capturing anomalous behavior,

4 Methodology

4.1 Phase 1: Anomaly-Based Risk Flagging (Isolation Forest)

To optimize fraud detection without increasing customer friction, the project implements an unsupervised anomaly detection layer. This phase serves as a non-intrusive "Flagging System" rather than a hard enforcement mechanism.

- **Strategic Flagging Logic:** Unlike traditional methods that immediately ban or block transactions, this system utilizes the *Isolation Forest* algorithm to assign a "Soft Flag." By analyzing key behavioral signals ($V3, V7, V10, V12, V14, V16, V17$),

it identifies patterns that deviate from standard consumer behavior without disrupting legitimate commerce.

- **Risk Score Calibration:** The model generates a standardized metric, `risk_score`. Empirical data analysis reveals a clear business threshold for risk prioritization:
  - **High-Risk Threshold (≥ 0.08):** Transactions exceeding this score exhibit a high probability of fraud and are prioritized for deeper analysis in the subsequent modeling stage.
  - **Normal Baseline:** Standard legitimate transactions typically cluster between -0.20 and -0.15, providing a stable baseline for "safe" behavior.

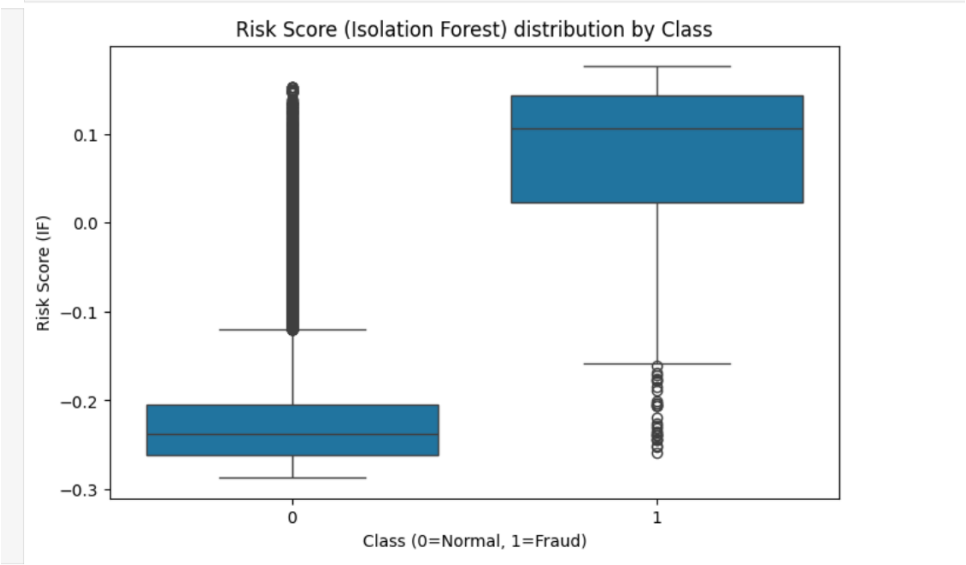

Figure 2: Risk Score from Train Data

-

- **Setting:** We set the the threshold is 0.05 as if risk score calculated is greater than this then the system will flag the transaction

- **Stakeholder and Operational Impact:** This strategy satisfies the **Risk Management** objective of high recall while fulfilling the Finance aspect while the later model does not need to classify all the data, hence reduces the operating cost.

.

4.2 Model

The modeling phase transitions from broad anomaly detection to a high-precision classification engine. **only transactions with a Risk Score exceeding the pre-defined**

threshold (≥ 0.05) are promoted to this stage for detection.

4.2.1 Evaluation Metrics

The detection challenge is framed as a binary classification task (Class 1: Fraud; Class 0: Legitimate). To align the system with banking priorities, the following metrics are used to measure success:

- **Recall for the Fraud Class:** proportion of true frauds correctly detected. Missing a fraud (false negative) can result in direct financial loss.
- **Precision for the Fraud Class:** proportion of predicted frauds that are truly fraudulent. Low precision implies many false positives, which can cause unnecessary customer friction and operational costs.
- **ROC AUC:** A global metric used to evaluate the system's ability to rank transactions by risk intensity across various decision thresholds.

4.3 Logistic Regression with Class Weighting

4.3.1 Performance Analysis

When evaluated on the time-based test set, the Logistic Regression strategy yields the following quantitative results:

- **Fraud Recall ($\approx$ 88%):** The model is highly aggressive in flagging potential fraud, successfully detecting nearly 88% of true fraudulent transactions.
- **Fraud Precision ($\approx$ 4%):** The precision for the fraud class is notably low, meaning that a significant number of legitimate transactions are flagged as suspicious (False Positives).
- **Overall ROC AUC ($\approx$ 0.95):** Indicates strong overall ranking quality across various thresholds.

4.3.2 Business Perspective

From a business operations standpoint, this model represents an aggressive security posture:

- The high recall effectively minimizes direct financial loss by capturing the vast majority of fraudulent strikes.

- However, the low precision introduces substantial "customer friction," as many legitimate users may face transaction delays or declines.
- **Operational Insight:** This configuration might be acceptable in a second-layer review process, where flagged transactions trigger a lightweight verification (such as an SMS confirmation). However, as a standalone front-line decision engine, this level of false positives may impose a heavy burden on operational teams and harm the customer experience.

4.4 XGBoost Gradient Boosting Model

4.4.1 Performance Analysis

To achieve a better balance between detection coverage and false alarms, the XGBoost strategy was implemented, yielding the following results on the temporal test set:

- **Fraud Recall ($\approx$ 78%):** While the recall decreases slightly compared to the baseline, the model still captures a robust majority of fraudulent activities.
- **Fraud Precision ($\approx$ 32%):** Precision increases substantially, meaning that a much larger portion of flagged transactions are truly fraudulent.
- **ROC AUC ($\approx$ 0.96):** Demonstrates superior discrimination power and high reliability in ranking transactions by risk level.

4.4.2 Business Perspective

Compared to the Logistic Regression model, this strategy represents a much more attractive trade-off for real-world banking operations:

- **Enhanced User Experience:** Fewer legitimate customers are disrupted, protecting transaction revenue and brand loyalty.
- **Resource Optimization:** The operational review workload is focused on a smaller, higher-quality set of alerts, allowing risk analysts to work more efficiently.
- **Model Selection:** Due to its superior balance between security and customer experience, the **XGBoost model is chosen as the primary engine** for deployment.

5 Deployment and User Interface

5.1 Streamlit Web Application

A lightweight Streamlit web application is developed as a demonstration interface for fraud risk scoring and decision support. The application is designed to reflect a realistic operational workflow in which anomaly detection is used as a preliminary risk filter before applying a supervised fraud classification model. The main capabilities of the application include:

- **File upload:** Users can upload a raw CSV file following the same schema as the original dataset, including `Time`, `V1--V28`, and `Amount`, with the `Class` label being optional.
- **Automatic preprocessing:**
  - Deduplicate rows and reset indices.
  - Fill missing numerical values with zero to ensure robustness.
  - Apply the saved `RobustScaler` to compute the scaled transaction amount.
  - Derive temporal features by extracting `hour_of_day` from `Time` and computing its sinusoidal representation.
  - Compute an unsupervised anomaly score (`risk_score_if`) using the trained Isolation Forest model.
  - Remove raw `Amount`, `Time`, and intermediate time-based features after feature engineering.
- **Risk-based filtering logic:** Transactions are first screened using the Isolation Forest risk score. Only records with `risk_score_if` > 0.05 are forwarded to the supervised XGBoost model for fraud probability estimation. Transactions with risk scores below this threshold are directly classified as non-fraud (label 0), reflecting a conservative and cost-aware screening strategy.
- **Fraud scoring:** For transactions that pass the risk filter, fraud probabilities are generated using the trained XGBoost model. These probabilities represent the likelihood of fraudulent behavior conditional on elevated anomaly risk.
- **Decision thresholding:** Fraud probabilities are converted into binary fraud flags based on a user-controlled decision threshold (default value of 0.5), allowing stakeholders to explore different trade-offs between recall and precision under constrained review capacity.

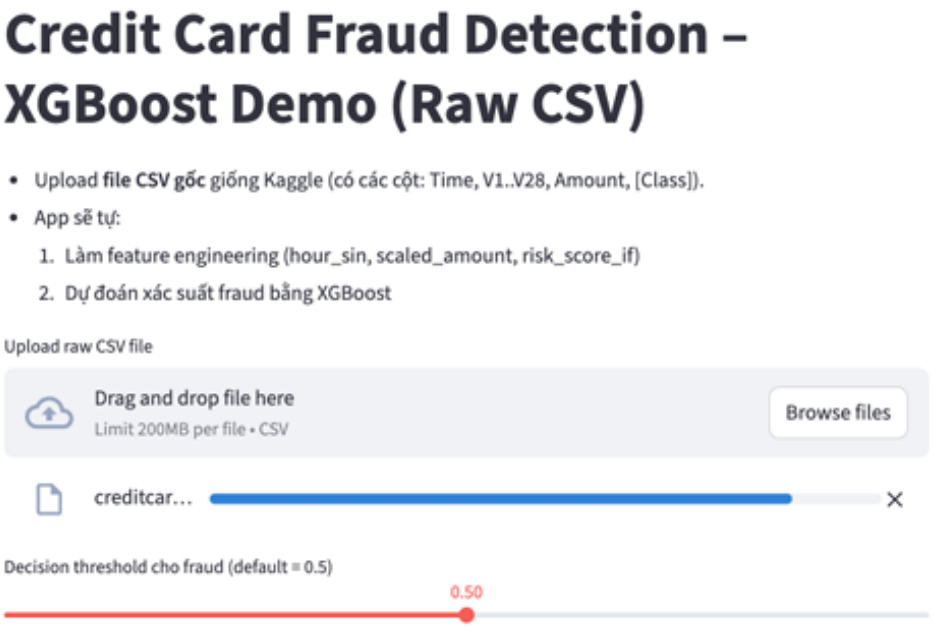

Figure 3: User Interface (UI) – Data Upload, Preprocessing, and Risk Filtering View

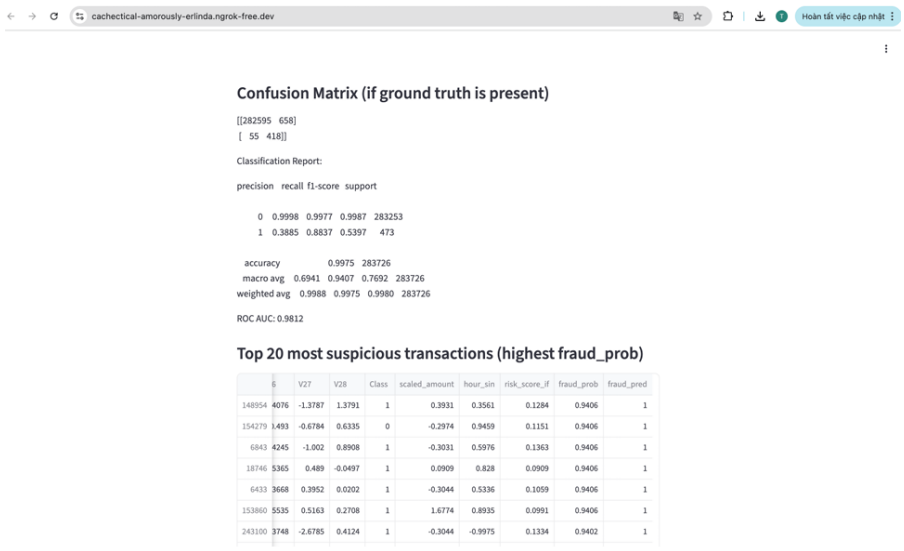

Figure 4: User Interface (UI) – Fraud Scoring, Threshold Analysis, and Results Visualization

6 Business Interpretation and Recommendations

6.1 Key Insights from the Data

From the exploratory analysis and modeling results, several meaningful business insights emerge:

- **Fraud is rare but impactful:** although only about 0.172% of transactions are labeled as fraud, the associated financial and reputational risks justify the use of advanced detection models.
- **Fraudsters favor inconspicuous amounts:** fraudulent transactions do not always correspond to extreme transaction amounts; many frauds use relatively small amounts to avoid triggering simple amount-based rules.
- **Time-of-day matters:** fraud occurrences increase during off-peak hours (approximately 1–5 AM), when normal customer activity is low. However, fraud is also present throughout the day, especially during business hours, likely to blend in with regular activity.
- **Anomaly-based risk scores are valuable:** the Isolation Forest-derived `risk_score_if` feature shows strong separation between fraud and non-fraud, providing an additional signal beyond raw features.

6.2 Operational Recommendations

Based on the final XGBoost model and its deployment, the following operational recommendations are proposed:

- **Implement a risk-based decision process:**
  - Use the model's fraud probability as a *risk score*.
  - Define multiple action bands:
    - *Very high risk:* automatically block transactions or require strong customer authentication (e.g., two-factor verification).
    - *Medium risk:* flag transactions for manual review or delayed settlement.
    - *Low risk:* allow transactions with standard monitoring.
- **Calibrate decision thresholds to business risk appetite:**
  - For higher protection, lower the decision threshold and accept more false positives.

  - For smoother customer experience, increase the threshold and reduce false positives, at the cost of missing some fraud cases.

- **Prioritize off-peak monitoring:** allocate additional automated or manual monitoring capacity to early-morning hours when fraud risk is relatively higher.

7 Limitations and Future Work

While the current solution demonstrates strong performance and a practical deployment path, several limitations and opportunities for improvement remain:

- **Limited feature set and anonymization:** due to confidentiality constraints, only PCA-transformed features are available. Access to more granular features (such as merchant category, geography, device identifiers, and customer history) could significantly enhance model performance and interpretability.

- **Short time horizon:** the dataset covers only two days of activity. Extending the analysis to longer time periods would enable more robust investigation of seasonal patterns, customer-level behavior, and model stability over time.

- **Cost-sensitive optimization:** while the current model is tuned using standard metrics (recall, precision, and ROC AUC), future work could incorporate explicit cost functions that reflect monetary losses from false negatives and operational costs associated with false positives.

- **Model explainability:** incorporating model explanation techniques (such as SHAP values) would help analysts understand which features drive individual fraud alerts, thereby improving trust and supporting policy and compliance decisions.

- **Continuous monitoring and retraining:** in a real production environment, the model should be periodically monitored for performance drift and retrained as fraud patterns and customer behavior evolve.

8 Conclusion

This project demonstrates how a combination of feature engineering, anomaly detection, and supervised learning can be used to build an effective credit card fraud detection system. By translating raw transaction data into actionable risk scores and deploying a usable interface, the solution provides tangible value to business stakeholders, including:

- improved detection of fraudulent transactions,
- more efficient allocation of investigation resources, and
- better control over the trade-off between customer experience and fraud losses.

Although the analysis is based on anonymized data with a limited temporal scope, the proposed methodology can be extended and adapted to richer, real-world environments. With further enhancements in feature design, cost-sensitive optimization, and model governance, the approach outlined in this work can serve as a solid foundation for an enterprise-grade fraud detection capability.