

Comment l'IA bloque une fraude bancaire en moins de 50 millisecondes

Machine learning, détection d'anomalies et apprentissage déséquilibré en temps réel

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques
tpineldahoui@gmail.com | github.com/Baudelaire12

09 Mars 2026

Série LinkedIn STEM au Quotidien — Jour 11/40

Plan de la présentation

- 1 Mise en contexte
- 2 Concepts clés
- 3 Développement mathématique
- 4 Cas pratique
- 5 Conclusion
- 6 Références

La fraude bancaire : un fléau mondial en accélération

Chiffre clé

En 2022, les pertes mondiales liées à la fraude par carte bancaire ont atteint **33,45 milliards de dollars**, soit une augmentation de 14% par rapport à 2021 (Nilson Report, 2023).

- Une transaction bancaire frauduleuse doit être détectée en **moins de 100 ms** pour être bloquée avant validation (VISA, 2022).
- Sur 1 million de transactions légitimes, on trouve en moyenne **100 à 500 fraudes** — soit un taux de fraude de 0,01 à 0,05% : le problème du déséquilibre de classes.
- En Afrique, la fraude mobile money est en forte croissance : les pertes estimées en 2022 dépassent **1 milliard USD** sur le continent (GSM Association, 2023).

Défi central

Comment détecter des transactions frauduleuses rarissimes parmi des millions de transactions légitimes, en temps réel et sans bloquer les clients légitimes ?

Le problème du déséquilibre de classes

Définition

Un jeu de données est dit **déséquilibré** (imbalanced) lorsque les classes cibles n'ont pas des effectifs comparables. En détection de fraude, les cas positifs (fraudes) représentent souvent moins de **0,1%** des données.

Conséquence naïve : un modèle qui prédit toujours « non-fraude » aurait une précision de 99,9% — mais détecterait **zéro fraude**.

Techniques de rééquilibrage :

Méthode	Type	Principe
SMOTE	Sur-échantillonnage	Génère des exemples synthétiques de fraude
ADASYN	Sur-échantillonnage	SMOTE adaptatif aux zones frontières
Random Undersampling	Sous-échantillonnage	Réduit la classe majoritaire
Class Weight	Pondération	Pénalise davantage les erreurs sur fraudes

SMOTE — Synthetic Minority Oversampling Technique

Algorithme SMOTE (Chawla et al., 2002) — génération d'exemples synthétiques :

$$x_{\text{new}} = x_i + \lambda \cdot (x_{k\text{-nn}} - x_i), \quad \lambda \sim \mathcal{U}[0, 1]$$

où x_i est un exemple de fraude, $x_{k\text{-nn}}$ l'un de ses k plus proches voisins dans la classe minoritaire.

Isolation Forest — Score d'anomalie :

Principe : les anomalies sont isolées plus rapidement dans un arbre de décision aléatoire.

$$s(x, n) = 2^{-\frac{E[h(x)]}{c(n)}}$$

où $h(x)$ est la longueur du chemin pour isoler x , et $c(n) = 2H(n-1) - \frac{2(n-1)}{n}$ la longueur de chemin moyenne dans un arbre de n nœuds (H : nombre harmonique).

Si $s(x, n) \rightarrow 1$: x est très probablement une anomalie (fraude).

Métriques adaptées aux données déséquilibrées

Ne pas utiliser l'exactitude (accuracy) en cas de déséquilibre !

Métriques pertinentes :

$$\text{Précision} = \frac{VP}{VP + FP} \quad (\text{des alertes levées, combien sont vraies ?})$$

$$\text{Rappel} = \frac{VP}{VP + FN} \quad (\text{des vraies fraudes, combien détectées ?})$$

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{Précision} \times \text{Rappel}}{\beta^2 \cdot \text{Précision} + \text{Rappel}}$$

En fraude, on choisit $\beta > 1$ pour privilégier le rappel (mieux vaut fausser une alerte que rater une fraude).

AUC-PR (aire sous la courbe Précision-Rappel) — plus robuste que l'AUC-ROC en contexte déséquilibré.

Dataset Credit Card Fraud — Kaggle/ULB

Jeu de données de référence

Dal Pozzolo et al. (2015) — Université Libre de Bruxelles — **284 807 transactions** réelles de 2 jours, dont seulement **492 fraudes** (0,172%). Données anonymisées par PCA.

Comparaison des modèles (avec SMOTE + optimisation seuil) :

Modèle	AUC-ROC	AUC-PR	Rappel	Précision
Régression logistique	0.974	0.712	0.88	0.63
Random Forest	0.985	0.854	0.90	0.89
XGBoost	0.991	0.891	0.93	0.91
Isolation Forest	0.967	0.731	0.82	0.71

Déploiement temps réel : Modèle XGBoost sérialisé + API REST → latence de décision < **50 ms** par transaction.

Synthèse

La détection de fraude bancaire est un cas d'école du machine learning en production : données massivement déséquilibrées, contrainte de latence extrême, et enjeux économiques énormes. XGBoost, les Forêts d'Isolation et les réseaux de neurones sont les algorithmes de référence du secteur.

Perspectives :

- Graph Neural Networks pour détecter les réseaux de fraude organisée (pas seulement les transactions isolées)
 - Federated Learning pour entraîner des modèles inter-banques sans partager les données clients
 - Adaptation aux spécificités africaines : fraude sur Mobile Money (M-Pesa, MTN MoMo, Wave)
 - Explicabilité (SHAP) pour justifier le blocage d'une transaction auprès des régulateurs
- « Les fraudeurs s'adaptent. Les algorithmes aussi. C'est une course permanente »*

Références bibliographiques

-  Nilson Report. (2023). *Card Fraud Losses Worldwide 2022*. The Nilson Report, Issue 1232.
-  Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE : Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
-  Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. *IEEE SSCI*, 159–166.
-  Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation forest. *IEEE ICDM*, 413–422.
-  Chen, T., & Guestrin, C. (2016). XGBoost : A scalable tree boosting system. *ACM SIGKDD*, 785–794.
-  GSM Association. (2023). *State of the Industry Report on Mobile Money*. GSMA.

Merci pour votre attention

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques

`tpineldahoui@gmail.com | github.com/Baudelaire12`

LinkedIn : @Pinel Baudelaire DAHOUI

Série STEM au Quotidien — 40 jours