

Projet REG08 : Credit Scoring ML

Prédire et expliquer le risque de défaut client

Intaset : Home Credit Default Risk



CHAKOURNE FREDERIC

ATSUVIA Marc Komi Christophe

BA Mohamed

BAH Mamoudou

Objectifs du Projet : Une Approche Systémique



Modélisation

Calculer une probabilité de défaut précise et robuste.



Décision

Transformer un score probabiliste en décision métier (**Accord/Refus**).



Interprétabilité

Expliquer les causes (Global & Local/SHAP) pour la transparence.



Interface

Dashboard interactif Streamlit pour les chargés de clientèle.



API

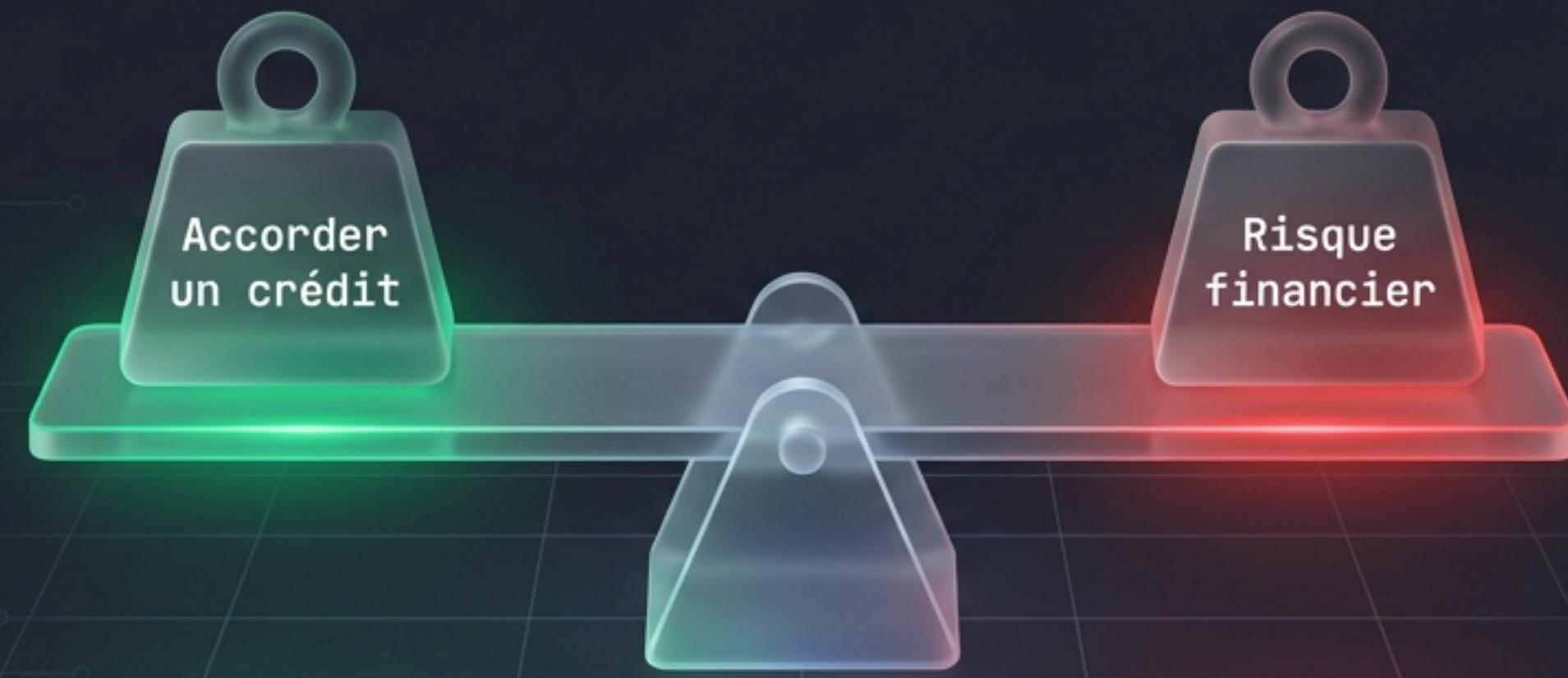
Architecture découpée pour des réponses temps réel.



Tracking

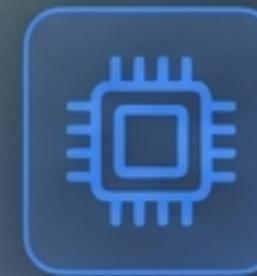
Suivi des expérimentations et reproduction via MLflow.

La Problématique Métier : Équilibre Risque vs Opportunité



Décision Client

Basée sur le profil financier et comportemental.



Automatisation

Traitement rapide et standardisé des demandes.



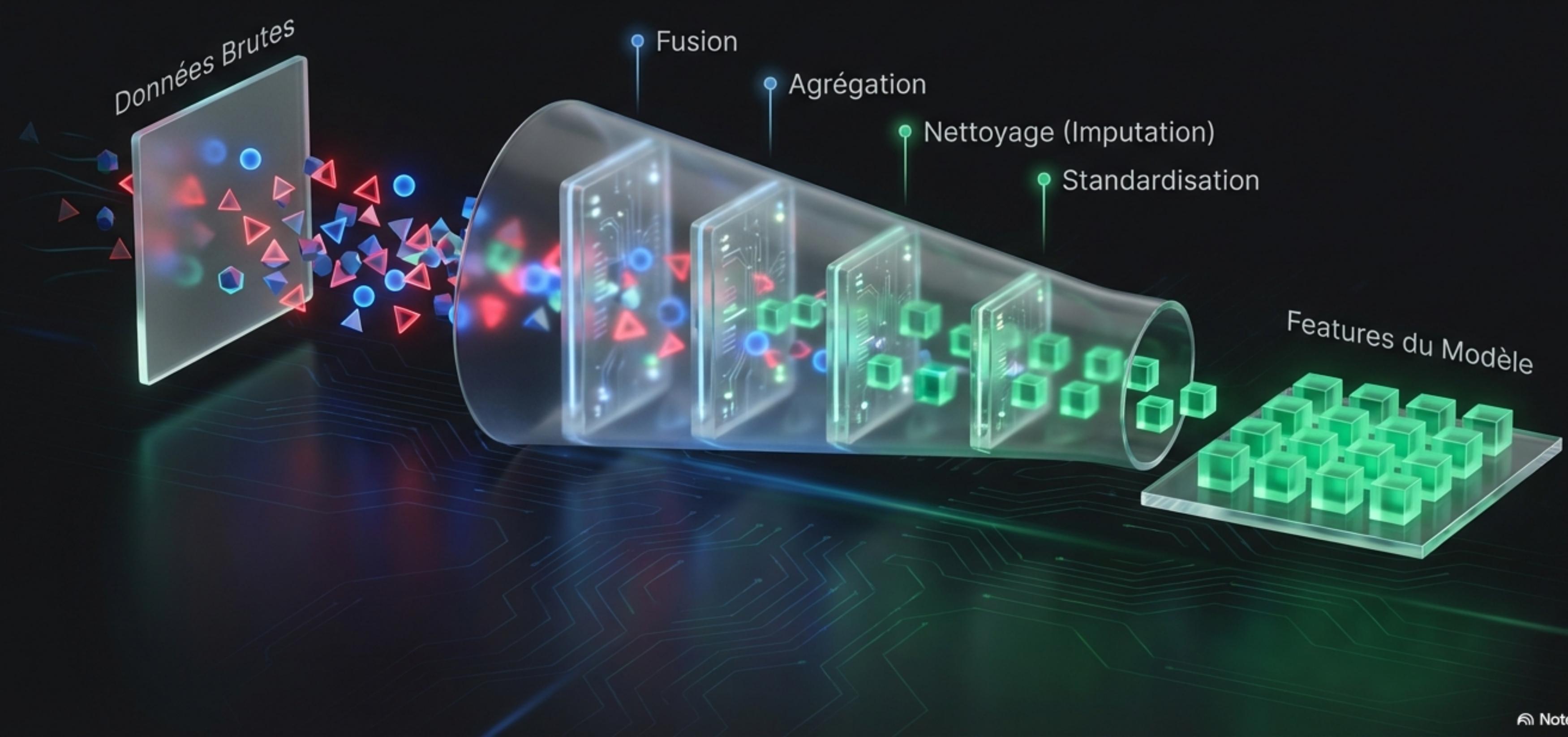
Réduction du Risque

Minimiser les défauts de paiement.

Écosystème de Données : Une Vue 360°



Préparation et Feature Engineering



Sélection du Modèle : La Stratégie du Champion

Régression Logistique

Baseline
Inter Regular

Random Forest

Modèle complexe
Inter Regular

Gradient Boosting

Champion (Performance Maximale)
Inter Regular

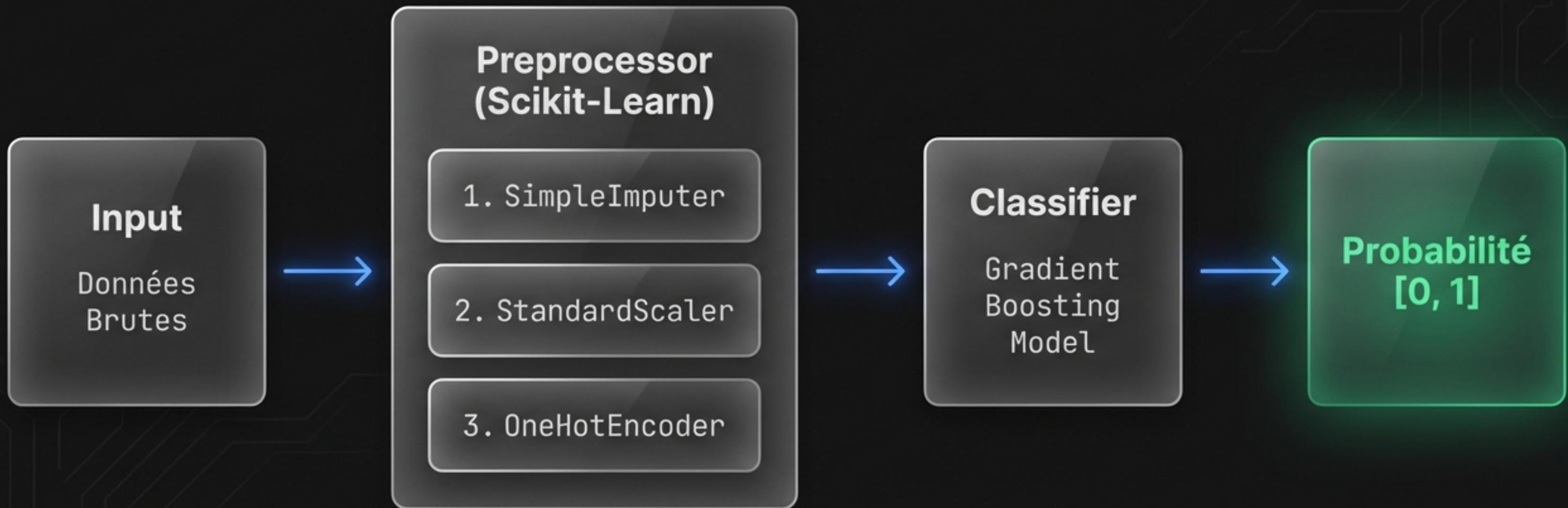
✓ Verified

Robustesse & Non-linéarité
Inter Regular

Performance : Analyse des Courbes ROC AUC



Pipeline Machine Learning : Flux Automatisé



Suivi des Expérimentations avec MLflow

1 Tracking

Run Name	Date	Parameters	Metrics (AUC)
Gradient Boosting	2023-10-27 10:00	{learning_rate: 0.1, n_estimators: 100, max_depth: 3}	0.757 Version Hash: abc123def456
Logistic Regression	2023-10-27 09:45	{C: 1.0, solver: lbfgs}	0.749
Random Forest	2023-10-27 09:30	{n_estimators: 50, max_depth: 5}	0.718

2 Métriques

Details

Artefacts

model.pkl

shap_plot.png

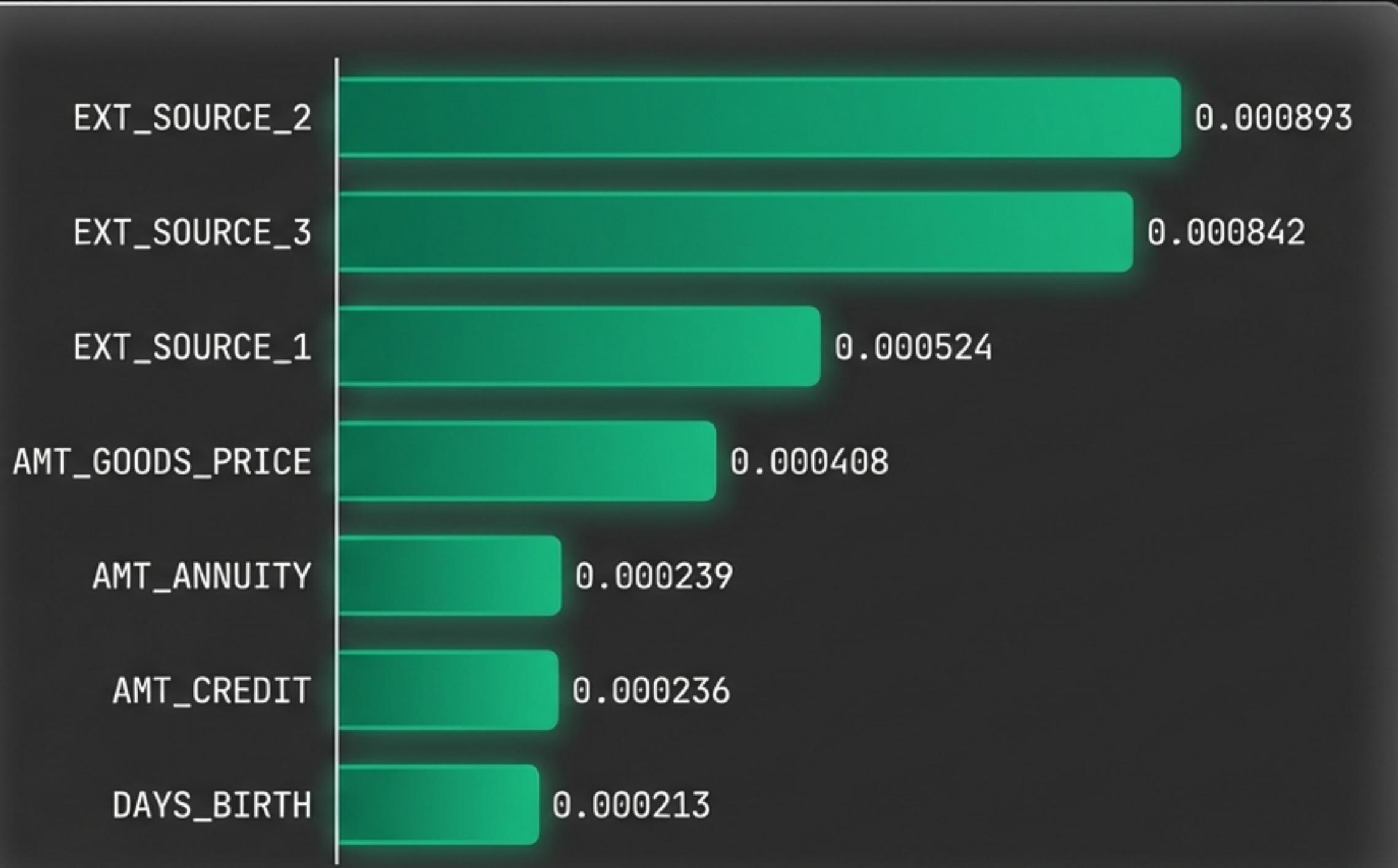
confusion_matrix.png

3 Artefacts

4 Traçabilité

Feature Importance (Globale)

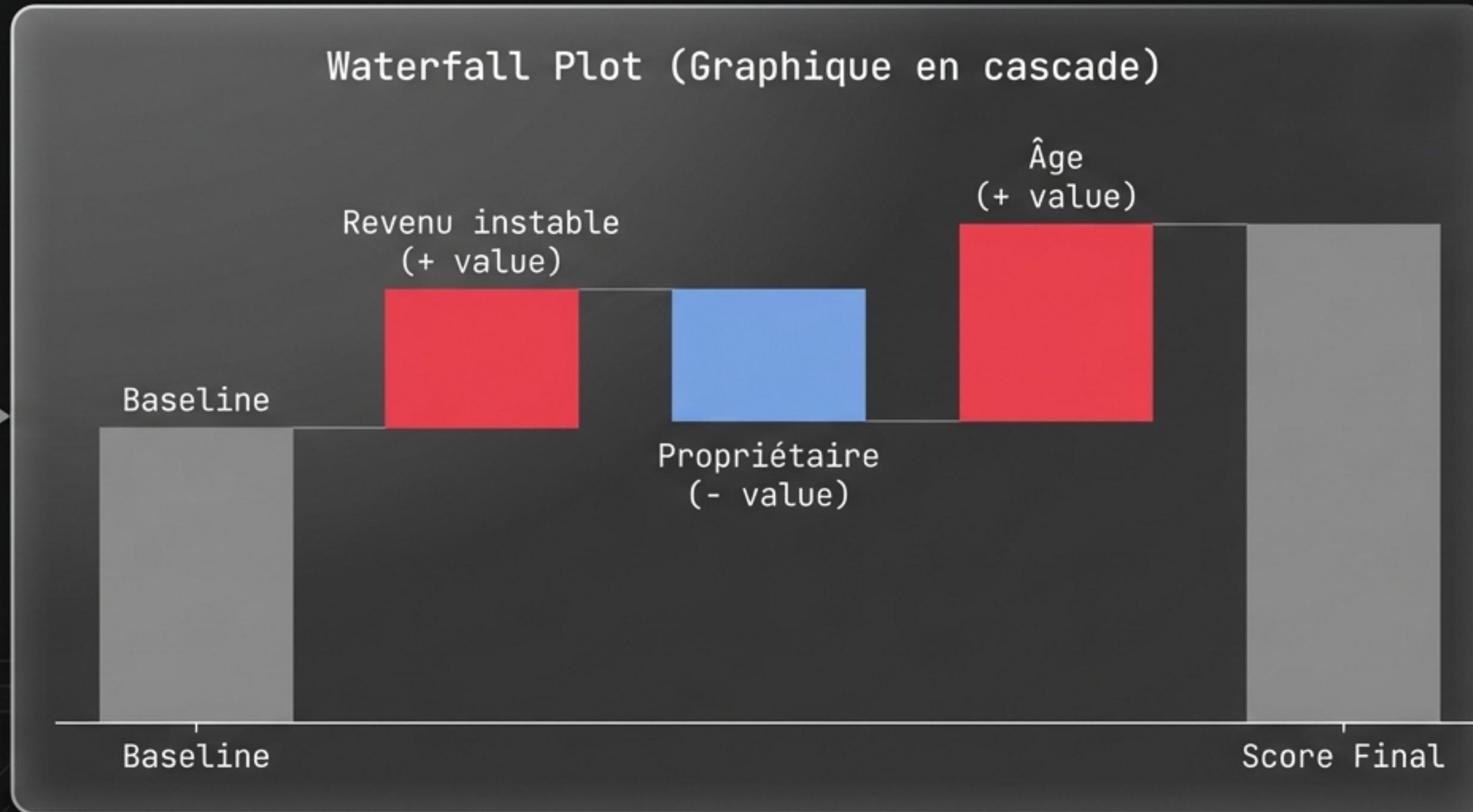
Les scores externes (EXT_SOURCE) et le montant du financement sont les leviers principaux du risque.



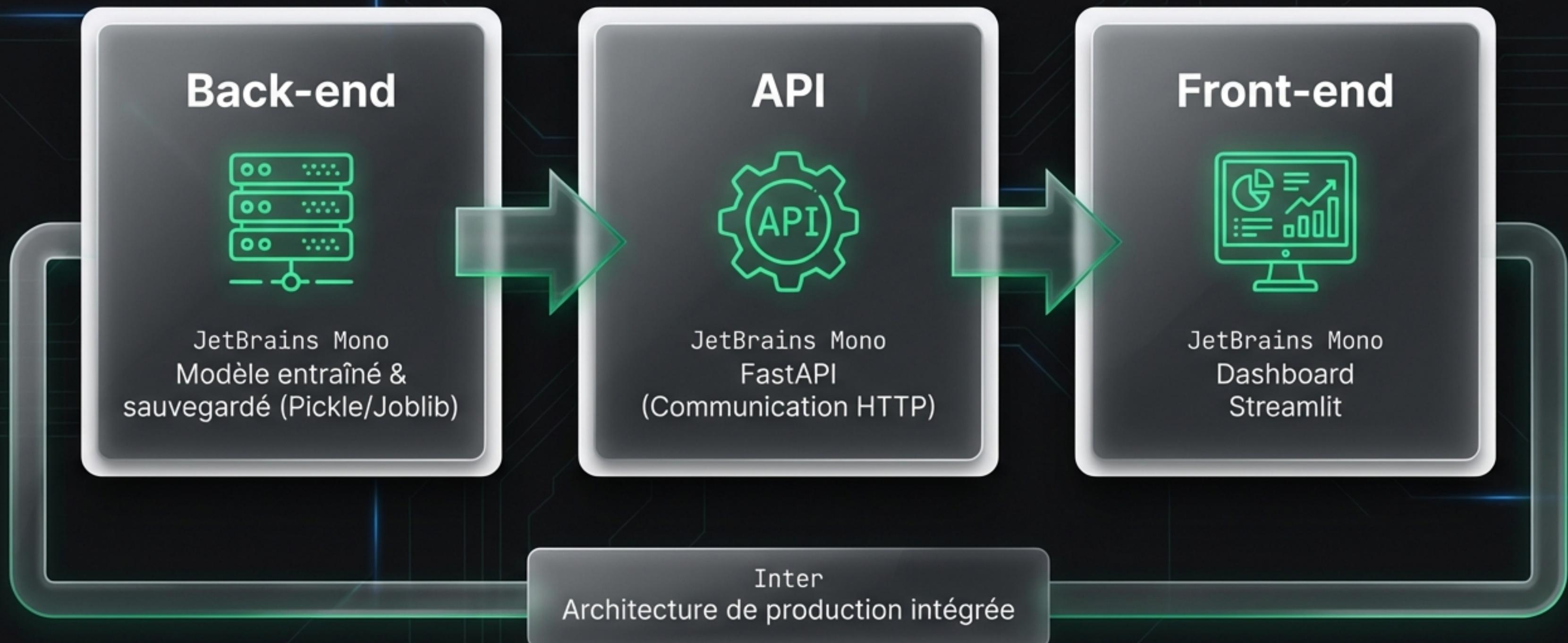
Transparence Décisionnelle : SHAP

Pourquoi ce client a-t-il ce score ?

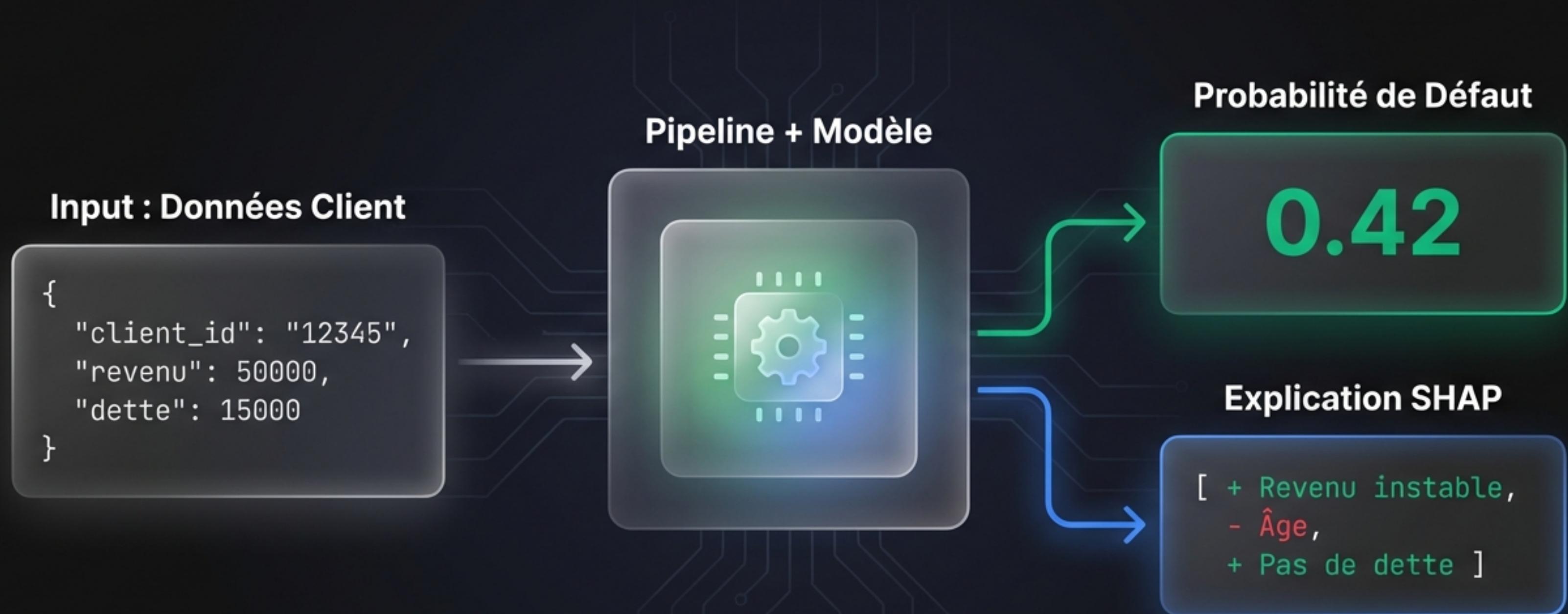
Explication
locale pour
chaque dossier



Architecture Système & MLOps



Le Moteur : API Credit Scoring



Traitement en temps réel pour une décision instantanée.

Logique de Décision et Interface



ACCORD

Seuil : 0.5

REFUS

Demande de crédit refusée

Risque estimé : 58.1%

Niveau de risque : Élevé

Pourquoi cette décision ?

- Fiabilité de paiement externe augmente le niveau de risque
- Historique financier externe augmente le niveau de risque
- Score de solvabilité externe augmente le niveau de risque

Bilan & Conclusion

✓ Points Forts

- Système complet de bout en bout (Data → Dashboard).
- Performance validée (AUC 0.757).
- Interprétabilité locale (SHAP) pour chaque client.

⚠ Limites & Vigilance

- Biais historiques potentiels dans les données.
- Sensibilité aux données manquantes.
- Nécessité de ré-entraînement régulier (Data Drift).