



Systeme de detection de Fraude

Identification en temps réel et prévention automatisée

Kévin Chatelain

Architecte en Intelligence Artificielle

Bloc 3 – « Analyse-predictive-de-
donnees-structurees-par-IA »

Jedha 2023/2024





Section 1 : Contexte et enjeux

Présentation des données, modèles et de l'API



Contexte et enjeux

Introduction générale

Problématique

Les risques liés aux fraudes dans les paiements en temps réel

Objectifs

- Détecter les paiements frauduleux avec précision.
- Fournir une infrastructure scalable et automatisée.

Valeur ajoutée

Prévention des pertes financières et amélioration de la sécurité.



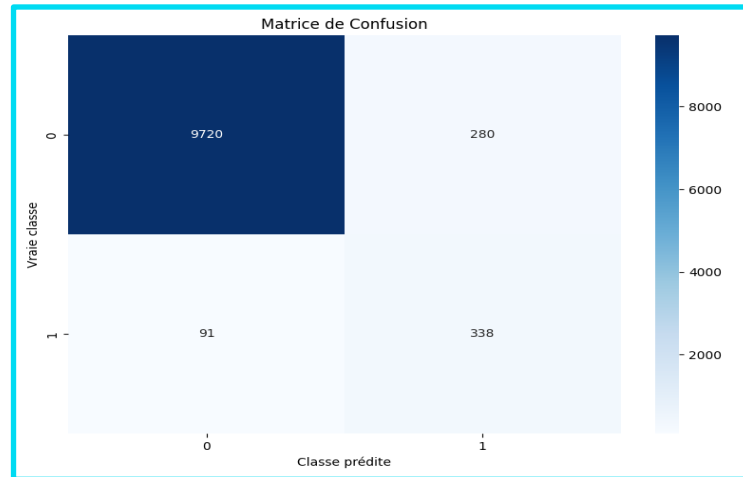
Contexte et ressources

Présentation du modèle

Random Forest (RF) : Performant pour les datasets déséquilibrés, robuste au bruit et adapté aux tâches de classification complexes.

Variable clés:

- Montant de la transaction.
- Distance géographique entre le marchand et le client.
- Horaire de la transaction.



Modèle	Précision (%)	F1-Score (%)	AUC-ROC (%)	Note du modèle
RF	97.5%	86.3%	94%	Meilleur modèle : haute précision et robustesse.



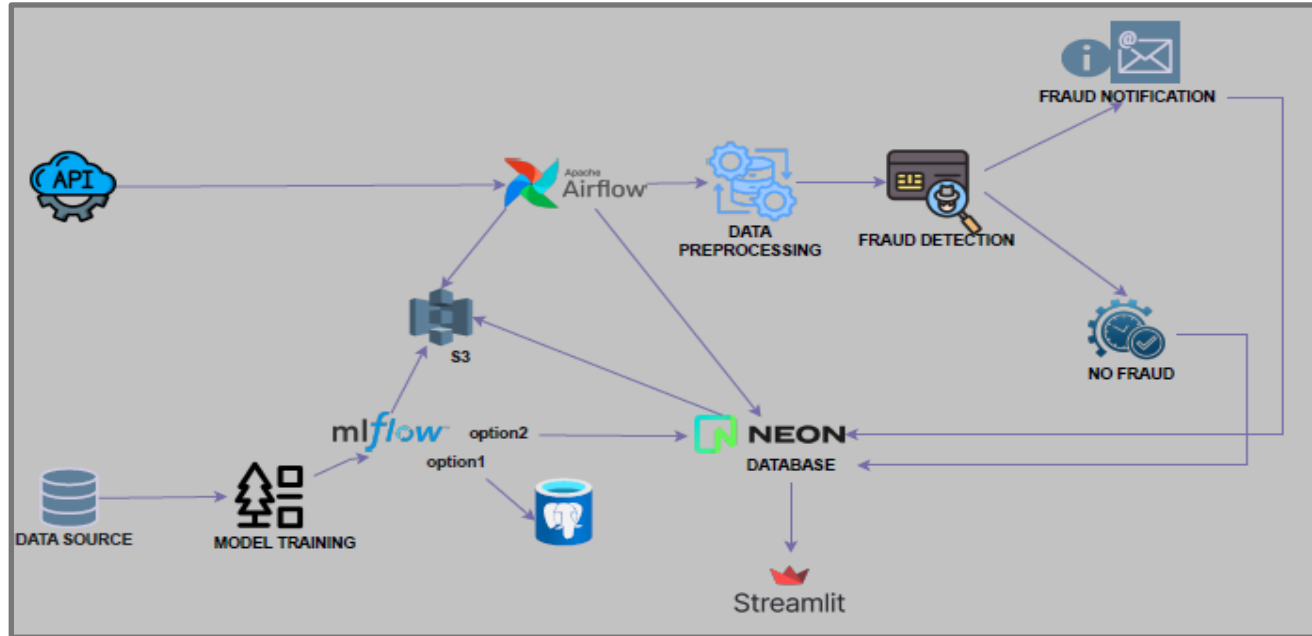
Section 2 : Architecture du projet

Présentation des données, modèles et de l'API



Architecture du projet

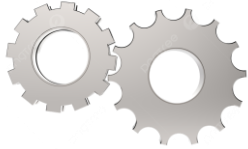
Schéma d'infrastructure





Architecture du projet

Avantages clés



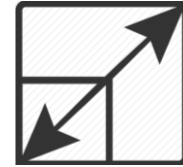
Automatisation

Processus entièrement automatisé



Temps réel

Analyse instantanée des transactions



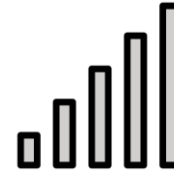
Scalabilité

Adapté aux volumes croissants



Traçabilité

Suivi complet des metrics



Interopérabilité

Intégration fluide des outils



Architecture du projet

Avantages clés



Monitoring robuste

Surveillance continue des systèmes



Sécurité

Sauvegarde centralisée et régulière



Accessibilité

Visualisation simplifiée



Optimisation des coûts

Utilisation efficace des outils cloud



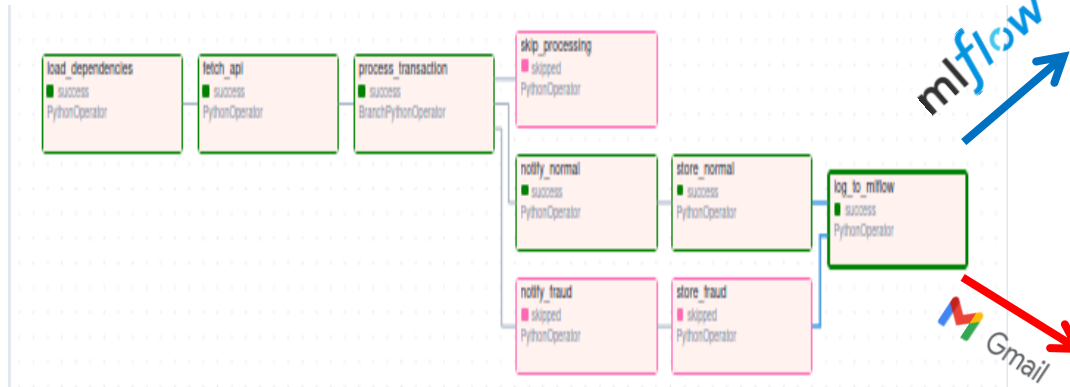
Section 3 : Pipeline de détection

Cheminement et outils



Pipeline de détection

Exécution avec Airflow et résultats



Run Name	Created	Dataset	Duration	Source	Models
fraud_detection_fixed_run	47 seconds ago	-	33.3s	airflow	-
fraud_detection_fixed_run	1 minute ago	-	30.7s	airflow	-
fraud_detection_fixed_run	2 minutes ago	-	28.1s	airflow	-
fraud_detection_fixed_run	2 minutes ago	-	36.6s	airflow	-
fraud_detection_fixed_run	3 minutes ago	-	48.2s	airflow	-

ALERTE : FRAUDE DETECTÉE Boîte de réception x

 **anneformation035@gmail.com**
A moi

Transaction FRAUDULEUSE détectée!

Probabilité de fraude: 76.47%

Détails de la transaction:

ID Transaction: 738cf6f98bf1f382e51e32224bfd7a5

Montant: \$284.07

Date/Heure: 2025-01-22 22:08:09

Informations sur le marchand:

Nom: fraud_Mosclski, Gislason and Mertz

Ville: Lorenzo

État: TX

Informations sur le client:

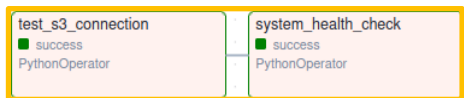
Nom: Christopher Adkins

Ville: Lorenzo



Pipeline de détection

Supervision et sauvegardes avec Airflow



Rapport quotidien - Système de détection de fraude Boîte de réception x

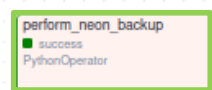
K chatelain.kev@gmail.com
À moi ▾

Rapport de santé du système:

API URL: <https://real-time-payments-api.herokuapp.com/current-transactions>
S3 Path: s3://fraud-detection-bucket-jedha/

Statut des services:

- api: ☒
- database: ☒
- storage: ☒



☒ Backup Neon DB - 202412 Boîte de réception x

K chatelain.kev@gmail.com
À moi ▾

Backup Neon DB effectué avec succès!

Résumé des données sauvegardées:

- fraud_transactions: 5 lignes
- normal_transactions: 155 lignes

Fichiers créés:

- s3://fraud-detection-bucket-jedha/backups/202412/fraud_transactions_202412.csv
- s3://fraud-detection-bucket-jedha/backups/202412/normal_transactions_202412.csv

Date: 2024-12-10 19:46:36



Pipeline de détection

Airflow -

Pipeline de détection

- Analyse en temps réel des transactions.
- Détection des fraudes avec notifications instantanées.
- Stockage des résultats pour analyse ultérieure.

Pipeline de sauvegarde

- Sauvegarde automatique des données critiques.
- Conservation historique des transactions.
- Prévention des pertes de données grâce à S3.

Pipeline de monitoring

- Vérification quotidienne de la santé des systèmes.
- Notifications en cas d'anomalies détectées.
- Garantie de la continuité des services.



Section 4: Données et visualisation

Outils de monitoring et de restitution



Données et visualisation

NeonDB

Stockage et gestion des métriques



- Haute disponibilité et scalabilité des données.
- Garantit la cohérence et le back-up des données.
- Requêtes SQL avancées pour des analyses complexes.

Suivi des expérimentations



- Centralisation des expérimentations (runs, métriques, artefacts).
- Visualisation claire des résultats de chaque itération.
- Reproductibilité des modèles grâce à la gestion des versions.

Visualisation interactive



- Tableaux de bord personnalisés pour visualiser les performances.
- Interactions dynamiques avec les données (filtres, graphiques).
- Accessibilité pour les utilisateurs non techniques.



Section 4: Axes d'améliorations

Optimisation & outils potentiels



Merci de votre Attention!

Dans l'attente de vos questions.

