Solution choisie : un fichier . json1

Où sont stockés les logs ?

Tous les logs sont enregistrés dans un fichier texte logs/predictions_log.jsonl, au format JSONL.

Chaque ligne est un objet JSON représentant un appel à l'API.

💡 Pourquoi le format JSONL ?

Le choix du format . j son1 repose sur plusieurs avantages clés pour ce projet :

- Format structuré : Chaque ligne est un objet JSON, donc très facile à parser en Python (via json.loads() ou pandas.read_json() avec lines=True).
- Scalabilité raisonnable : Contrairement au JSON classique, les logs peuvent être lus ligne par ligne sans devoir charger tout le fichier en mémoire.
- Exploitable pour l'analyse et le retraining : Ce format est parfaitement compatible avec les outils comme Pandas, Evidently, ou même un futur pipeline de réentraînement automatique.
- Pas besoin de base de données : Permet un stockage persistant sans complexité technique supplémentaire (pas de serveur PostgreSQL, MongoDB à maintenir).
- **Facile à versionner** : Le fichier . log peut être suivi via Git (ou stocké ailleurs), utile pour garder un historique.

Chaque log comprend:

- (E) timestamp : la date/heure UTC de la prédiction
- input : les données fournies par l'utilisateur
- @ prediction : score de risque + étiquette (faible/élevé)
- 🐧 duration : temps d'exécution de l'appel
- iii cpu_percent et memory_usage_MB : utilisation système

Utilisation des logs pour l'analyse

Ces logs sont utilisés dans plusieurs scripts :

- analyse_logs.py : calcule les temps de réponse, taux d'erreur, moyenne CPU, etc.
- drift_report.html (via evidently): détecte les dérives sur les features d'entrée
- dashboard.py: visualisation en Streamlit ou Gradio

