



Soirée Recommandation — Application IA

Projet présenté par **Nathan Serfaty**

Albert School, 2025

Introduction



Objectif

Recommander automatiquement des lieux de sortie selon le profil utilisateur.



Motivation

Créer une application simple inspirée de Netflix et Spotify pour les sorties.



Enjeu principal

Formuler des recommandations crédibles avec des données fictives.



Notre Dataset

Caractéristiques

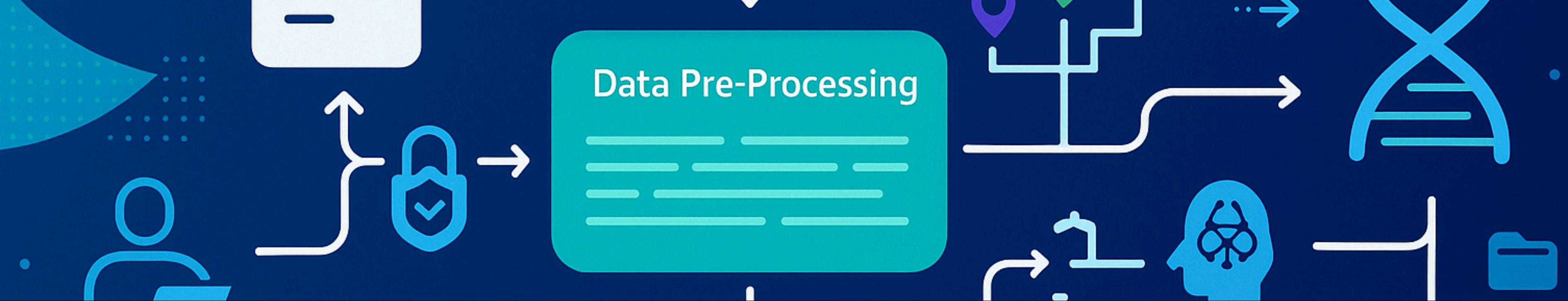
500 000 lignes générées avec Faker

Données 100% synthétiques

But: prédire le type de lieu selon les préférences

Variables clés

- Âge (entier)
- Genre (homme, femme, non-binaire)
- Ville (villes françaises)
- Préférence (type souhaité)
- Type (catégorie réelle)
- Nom du lieu (fictif)
- Note moyenne (1 à 5)



Prétraitement des Données

Nettoyage

Élimination des valeurs manquantes avec dropna.

Encodage

LabelEncoder pour la classification et OneHotEncoder pour la similarité.

Préparation

Séparation X/y pour l'entraînement des modèles.

Sauvegarde

Conservation des modèles et encodeurs avec joblib.

Structure du code

Architecture Générale

- app/
- **main.py**
- **model.py**
- **similarity.py**
- **preprocess.py**
- **templates/index.html**

Données & Résultats

- **data/dataset_sorties_500k.csv**
- **model/model.joblib**
- **model/metrics.json**
- **model/resultats_du_jour.json**
- **model/similarity_encoded_array.j
oblib**
- **static/recommandations_du_jour
.csv**

Entraînement & Configuration

- **train_model.py**
- **train_similarity.py**
- **requirements.txt**
- **README.md**
- **.gitignore**

Méthodes

Classification

- Modèles: XGBoost, Random Forest
- Prédit le type de lieu
- Rapide et interprétable
- Manque de nuance

Recommandation par

- Matching personnalisé
- OneHotEncoding des colonnes
- Calcul de cosine_similarity
- Approche plus flexible



Résultats

Modèle	Accuracy	F1-score	Précision	Rappel
XGBoost	0.92	0.91	0.90	0.92
Random Forest	0.89	0.88	0.87	0.89

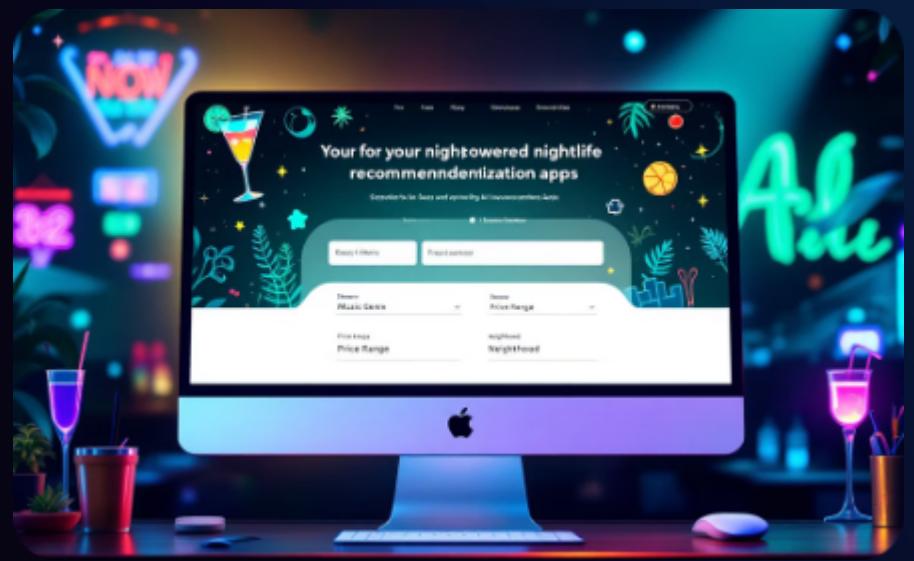
XGBoost a été sélectionné comme modèle final pour ses performances supérieures.

Fonctionnalités ajoutées: téléchargement des recommandations et export JSON.

XGBoost vs Random forest performances:

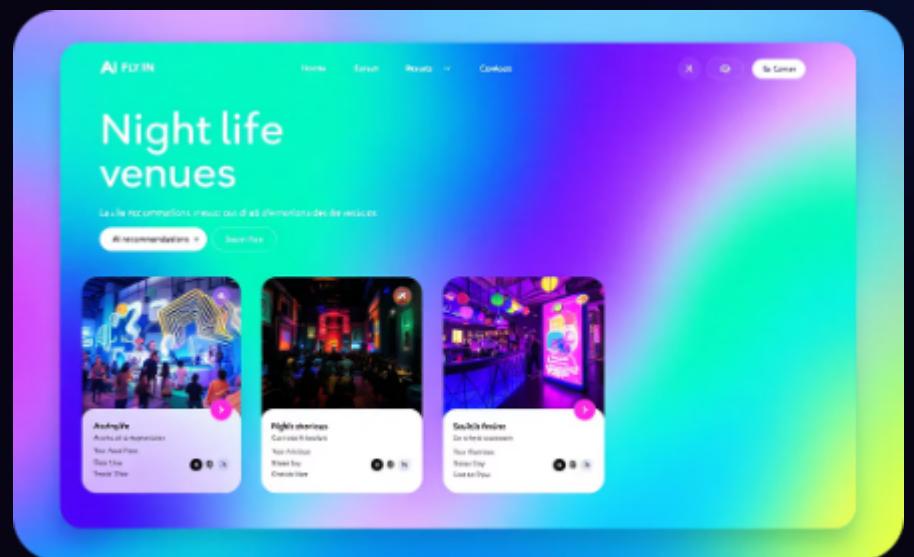


L'interface



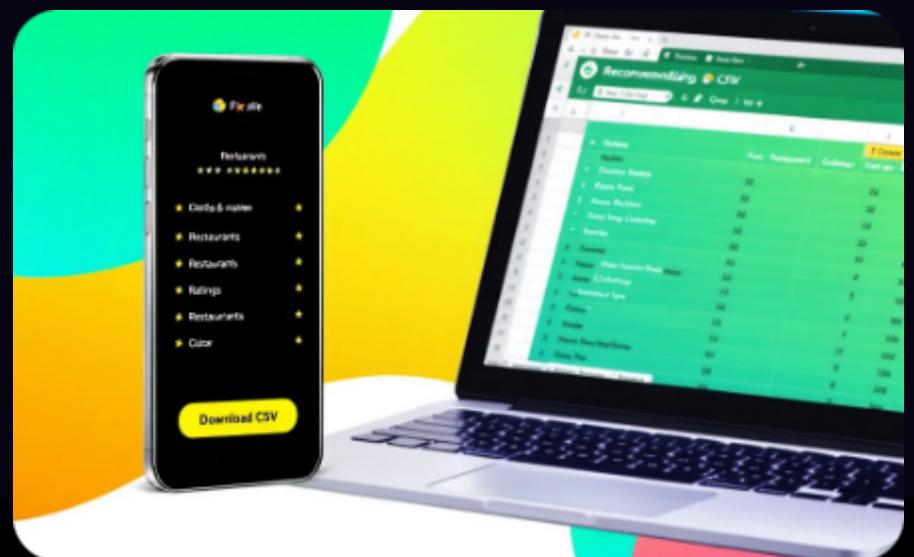
Interface Flask

Application web simple et intuitive.



Deux Formulaires

Recommandation générale et précise.



Exportation

Téléchargement CSV et consultation JSON.

Remarques Finales

Données simulées

Pipeline réaliste malgré l'utilisation de données fictives.



Projet extensible

Possibilité d'intégrer des données réelles (Yelp/TripAdvisor).

Améliorations

Google Maps API, suivi des préférences, moteur FAISS.

Démonstration Live

