
Cécile Guillot

Ingénieur Machine Learning
cecile.gltslmcs@protonmail.com

Algorithmes de recommandations appliqués aux livres et émissions de CO2

Janvier 2022

VUE D'ENSEMBLE

Le but de ce projet est de montrer la faisabilité de la création d'un algorithme de recommandation pour des livres. On trouve énormément de recherches sur les moteurs de recommandations de films, de séries ou de musiques mais très peu concernant les livres. L'idée est de voir si les données dont l'on dispose peuvent permettre d'appliquer les algorithmes de recommandations pour des livres sans que cela ne soit appliqué à une plateforme d'e-commerce.

OBJECTIFS

1. Tester la faisabilité de la création d'un algorithme de recommandations pour les livres
2. Mesurer les émissions de CO2 en lien avec les différentes approches

GRANDES ÉTAPES

Acquisition des données

- Trouver un jeu de données permettant l'implémentation d'algorithmes de recommandations
- Choisir entre le dataset GoodReads et Bookcrossing

Analyse des données

- Analyser le contenu des jeux de données pour voir les algorithmes à appliquer

-
- Préparer les données pour la comparaison des algorithmes

Baseline

- Déterminer les performances baseline si possible
- Déterminer une méthodologie à suivre (Bhowmick et al., 2021)

Modélisation

- Quel modèle pour quelle donnée ?
- Tests de modèles “user-based”
- Tests de modèles “content-based” avec utilisation de transformers
- Tests de modèles hybrides (si possible)