

Projet : systèmes de recommandation sur MovieLens

Kpodjro KPATOUKPA Serigne Bassirou NDIAYE
Lyse Priscille NGASSEU NDIFO
Orchelle Patricia WELEHELA TAWEUTEU

Janvier 2026

Université Paris Cité

1 Présentation générale

Ce projet s'inscrit dans le cadre du cours *Introduction to Recommendation Systems* et vise à mettre en pratique les notions fondamentales des systèmes de recommandation à travers l'implémentation et la comparaison de plusieurs modèles appliqués au jeu de données **MovieLens**. L'objectif est d'explorer différentes manières de modéliser les préférences des utilisateurs et d'évaluer la qualité des recommandations produites.

La problématique centrale est la suivante : *comment recommander efficacement des films à un utilisateur en exploitant à la fois ses interactions passées et les informations textuelles disponibles sur les contenus ?*

Pour répondre à cette question, trois approches ont été étudiées :

- **Matrix Factorization (MF)**, utilisée comme modèle collaboratif de référence.
- Un modèle *content-based* basé sur les **embeddings Qwen** (*Qwen2.5-0.5B-Instruct*), permettant de représenter les films à partir de leurs descriptions textuelles.
- Un modèle *content-based* basé sur **BERT** (*all-MiniLM-L6-v2*), exploré comme alternative afin d'évaluer l'impact de différentes représentations textuelles sur la qualité des recommandations.

Ces modèles sont évalués selon les métriques, **Recall@10** et **NDCG@10**, afin de comparer leur capacité à retrouver et à classer les films pertinents pour chaque utilisateur.

2 Données

Le projet utilise le jeu de données **MovieLens 100K**, qui contient des interactions explicites entre utilisateurs et films (notes) chargées via la bibliothèque **Cornac**, ainsi que des informations additionnelles sur les *items*, notamment les descriptions textuelles récupérées séparément depuis le site officiel de MovieLens. Il comporte **943 utilisateurs** et **1682 films**. Ces descriptions ont été nettoyées puis encodées par Qwen et BERT afin d'obtenir des embeddings permettant de représenter les films dans un espace vectoriel sémantique et de calculer des similarités entre contenus.

Pour garantir une évaluation cohérente entre les différents modèles, les données ont été prétraitées (voir notebook) et séparées en deux ensembles :

- **Ensemble d'entraînement** : utilisé pour apprendre les représentations collaboratives (MF) et pour construire les profils utilisateurs dans les modèles *content-based*.
- **Ensemble de test** : utilisé pour mesurer la capacité des modèles à recommander des films non vus, selon les métriques **Recall@10** et **NDCG@10**.

Les recommandations **Top-10** sont obtenues en classant les items non vus selon leur score prédictif (Factorisation matricielle) ou leur similarité cosinus avec le profil utilisateur (Qwen, BERT).

Un film est considéré comme pertinent s’il a reçu **une note** ≥ 4 de l'utilisateur dans l'ensemble de test ; le Top-10 est ensuite évalué en vérifiant combien de ces films pertinents apparaissent parmi les dix recommandations proposées.

3 Résultats

Les performances des trois modèles implémentés sont évaluées sur l'ensemble de test selon les métriques **Recall@10** et **NDCG@10**. Les résultats obtenus sont présentés dans le tableau suivant :

Table 1: Performances des modèles de recommandation

Modèle	Recall@10	NDCG@10	Type d'approche
MF	0.0184	0.0122	collaboratif (latent)
Qwen	0.0155	0.0201	contenu (LLM)
BERT	0.0254	0.0398	contenu (sentence)

De manière générale, la factorisation matricielle est souvent considérée comme une approche efficace pour retrouver les items pertinents, grâce à l'exploitation des interactions collaboratives. Dans notre cas toutefois, BERT se distingue nettement avec des performances supérieures sur les deux métriques, notamment un Recall@10 plus élevé et un NDCG@10 presque doublé par rapport à Qwen.

Cette comparaison met en évidence l'impact déterminant du choix des embeddings textuels sur la qualité des recommandations.

4 Conclusion

Il était question ici d'évaluer différentes approches de recommandation en comparant un modèle collaboratif classique (Matrix Factorization) à deux variantes *content-based* reposant sur des embeddings de langage (Qwen et BERT). Il en ressort que les embeddings BERT offrent un compromis intéressant entre simplicité (modèle pré-entraîné, léger (MiniLM), facile à charger) et efficacité.

Force est de constater que la nature des représentations textuelles ont un impact direct sur la qualité des recommandations : dans notre cas, Qwen demeure moins performant pour cette tâche spécifique, ses embeddings plus généraux étant moins adaptés à la capture de similarités fines entre synopsis contrairement à BERT, qui est spécifiquement entraîné pour la similarité de phrases.

Au vu de tout cela, ce travail ouvre la voie à des approches hybrides combinant signaux collaboratifs et information textuelle, qui pourraient améliorer la qualité globale des recommandations. L'ensemble du code et du protocole expérimental est disponible dans le dépôt GitHub associé au projet suivant : <https://github.com/LyseNdifo/recommandation-systems>.