



Systeme de Recommandation des Produits

Réalisé par :

EDDAHBY Mohamed

Responsable :

Prof. ZEROUAL Imad





Plan

01

Introduction

02

**Collection et
Exploration des
Données**

03

**Prétraitement des
Données**

04

Méthodologie

05

Évaluation

06

Conclusion

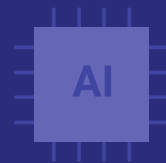




01

Introduction

- Importance des systèmes de recommandation
- Objectifs du projet



Introduction

Importance des systèmes de recommandation

Dans le monde dynamique du commerce électronique, la découverte rapide et précise de produits pertinents est essentielle. C'est là que les systèmes de recommandation entrent en jeu, en simplifiant le processus d'achat et en améliorant l'expérience utilisateur.

Les objectifs du projet :

Les objectifs clés de notre initiative incluent une exploration approfondie des données, la mise en œuvre de différentes approches de recommandation, et l'évaluation de leur performance respective.



Introduction

Challenge : Problème du Démarrage à Froid (Cold Start Problem)

Il est crucial de reconnaître l'un des défis majeurs auxquels ces systèmes sont confrontés : le "**Problème du Démarrage à Froid**" (**Cold Start Problem**).

Les systèmes de recommandation font face à des difficultés lorsqu'ils sont confrontés à de nouveaux utilisateurs ou de nouveaux produits pour lesquels ils disposent de peu ou pas d'historique d'interaction. Cette situation peut entraver la capacité du système à fournir des recommandations précises et personnalisées.

Introduction

Solutions Envisagées :

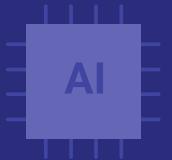
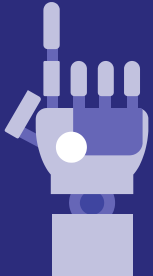
Pour surmonter le Problème du Démarrage à Froid, notre méthodologie intègre des approches telles que les recommandations basées sur le classement, le filtrage basé sur le contenu, et des solutions hybrides. Ces stratégies permettent d'offrir des recommandations pertinentes même en l'absence d'informations préalables significatives.



02

Collection et Exploration des Données

- Données utilisées
- Analyse Exploratoire



Collection et Exploration des Données

Jeu de données utilisé :

Nous avons utilisé un jeu de données d'Amazon sur les évaluations des utilisateurs pour les produits électroniques. Cet ensemble de données ne contient pas d'en-têtes. Afin d'éviter les biais, chaque produit et utilisateur est attribué un identifiant unique au lieu d'utiliser leur nom ou toute autre information potentiellement biaisée.

Lien :

<https://www.kaggle.com/datasets/vibivij/amazon-electronics-rating-datasetrecommndation/download?datasetVersionNumber=1>



Collection et Exploration des Données

Analyse Exploratoire :

Nous avons effectué une analyse approfondie, examinant la forme des données, les data types, les valeurs manquantes, et en fournissant des statistiques résumées pour mieux comprendre notre ensemble de données.

Nous avons identifié le nombre d'utilisateurs uniques, le nombre de produits uniques, et avons mis en lumière les utilisateurs les plus actifs dans nos données.

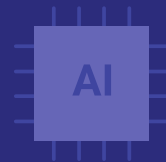
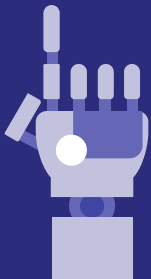
	user_id	prod_id	rating
0	AKM1MP6P0OYPR	0132793040	5.0
1	A2CX7LUOHB2NDG	0321732944	5.0
2	A2NWSAGRHCP8N5	0439886341	1.0
3	A2WNBOD3WNDNKT	0439886341	3.0
4	A1GI0U4ZRJA8WN	0439886341	1.0



03

Prétraitement des Données

- Optimisation des Données pour des Recommandations Puissantes



Prétraitement des Données

Sélection des Données :

Création d'un sous-ensemble en ne conservant que les utilisateurs actifs.

Matrice d'Interaction :

Transformation des données en une matrice d'interaction utilisateur-produit.

Densité des Avis :

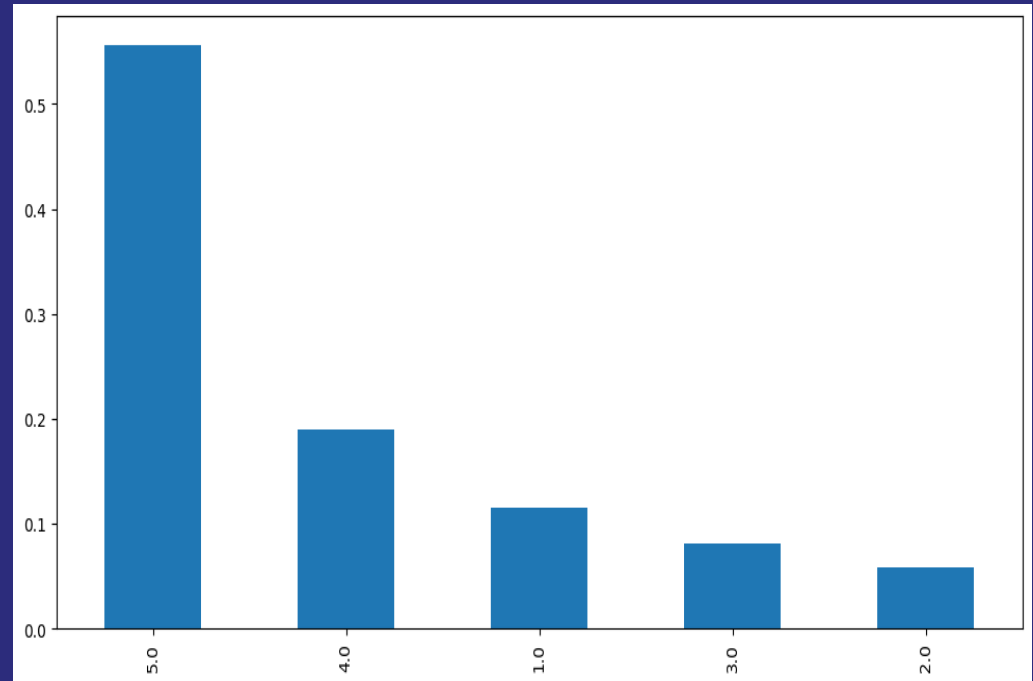
- Évaluation de la densité de la matrice d'interaction.
- Préparation des données pour des recommandations plus efficaces.



Prétraitement des Données

La distribution est décalée vers la droite.
Plus de 50 % des avis sont des avis 5 étoiles, suivis d'un peu moins de 20 % d'avis 4 étoiles.

Les pourcentages d'évaluation continuent de diminuer jusqu'à ce que moins de 10 % des évaluations soient des évaluations à 2 étoiles.

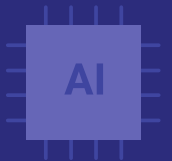


Distribution après prétraitement

04

Méthodologie

- Recommandation Basée sur le Classement
- Filtrage Collaborative Basé sur la Similarité
- Filtrage Collaborative Basée sur le Modèle (SVD)



Recommandation Basée sur le Classement

Objectif :

- Recommander des produits avec le plus grand nombre d'avis.
- Cibler de nouveaux clients avec les produits les plus populaires.

Fonctionnement :

1. Calcul des moyennes et du nombre total d'avis pour chaque produit.
2. Classement des produits en fonction de leur moyenne.
3. Recommandation des produits les mieux classés.



Recommandation Basée sur le Classement

Avantages :

- Simple et efficace.
- Résout le problème du démarrage à froid.

Résultats :

- Recommandation des 5 meilleurs produits avec un minimum de 50 interactions en fonction de leur popularité

✓
0s

```
[31] list(top_n_products(final_rating, 5, 50))
```

```
['B001TH7GUU', 'B003ES5ZUU', 'B0019EHU8G', 'B006W8U2MU', 'B000QUUFRW']
```

- Recommandation des 5 meilleurs produits avec un minimum de 100 interactions en fonction de leur popularité

✓
0s

```
[32] list(top_n_products(final_rating, 5, 100))
```

```
['B003ES5ZUU', 'B000N99BBC', 'B007WTAJTO', 'B002V88HFE', 'B004CLYEDC']
```

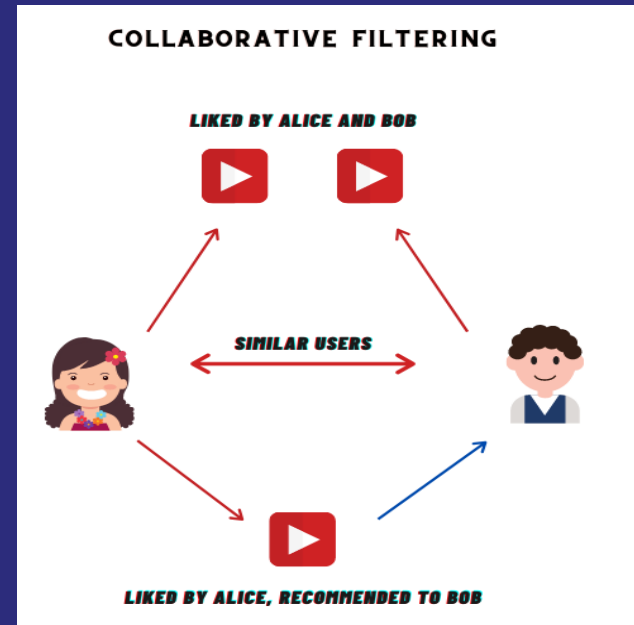
Filtrage Collaboratif Basé sur la Similarité

Objectif :

- Fournir des recommandations personnalisées et pertinentes aux utilisateurs.

Fonctionnement :

- Calcul de la similarité cosinus entre utilisateurs.
- Identification des utilisateurs similaires.
- Recommandation de produits appréciés par des utilisateurs similaires.



Filtrage Collaboratif Basé sur la Similarité

Avantages :

- Personnalisation des recommandations.
- Adapté à la diversité des goûts.

Résultats :

- Recommandation de 5 produits à l'index 4 de l'utilisateur sur la base d'un filtrage collaboratif basé sur la similarité.

```
[44] recommendations(4,5,final_ratings_matrix)
```

```
['B009X00E18', 'B00FPG37R2', 'B002VPE1X4', 'B00ATM1MVU', 'B005I6EU48']
```

- Recommandation de 5 produits à l'index 1000 de l'utilisateur sur la base d'un filtrage collaboratif basé sur la similarité.

```
[45] recommendations(1000,5,final_ratings_matrix)
```

```
['B004GUT1Q0', 'B00BQHD4B8', 'B002IY35WQ', 'B0010I2Z4Q', 'B00028DAKC']
```

Filtrage Collaboratif Basé sur le Modèle

Objectif :

- Fournir des recommandations personnalisées aux utilisateurs basées sur leur comportement et leurs préférences passés, tout en traitant les défis de la sporadicité et de la scalabilité qui peuvent survenir dans d'autres techniques de filtrage collaboratif.

Fonctionnement :

1. Décomposition en valeurs singulières (SVD) de la matrice d'interaction utilisateur-produit.
2. Prédiction des évaluations pour tous les utilisateurs.
3. Recommandation des produits avec les évaluations prédites les plus élevées.



Recommandation Basée sur le Classement

Avantages :

- Gestion efficace de la sporadicité.
- Recommandations plus précises grâce aux caractéristiques latentes.
- Recommandation personnalisée pour chaque utilisateur.
- Adapté aux jeux de données volumineux et clairsemés.

Résultats :

- Recommandation des 5 meilleurs produits à l'utilisateur id 100

```
[55] recommend_items(100,final_ratings_sparse,preds_matrix,5)
```

Below are the recommended products for user(user_id = 100):

Recommended Products

11078	1.624746
16159	1.132730
10276	1.047888
22210	0.955049
18887	0.879705

Name: user_predictions, dtype: float64

- Recommandation des 10 meilleurs produits à l'utilisateur id 150

```
[56] recommend_items(150,final_ratings_sparse,preds_matrix,10)
```

Below are the recommended products for user(user_id = 150):

Recommended Products

38250	1.006201
33113	0.991509
24827	0.933169
22460	0.860471
27408	0.713468
38655	0.602994
38615	0.580360
18510	0.535426
11870	0.533274
31355	0.523534

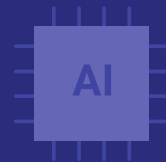
Name: user_predictions, dtype: float64



05

Évaluation

- Mesurer la Performance du system de recommandation



Évaluation

Évaluation du Modèle SVD :

RMSE (Root Mean Squared Error) : 0.0137

Le RMSE mesure la précision du modèle en termes de différence entre les évaluations réelles et les évaluations prédites. Un RMSE plus bas indique une meilleure adéquation.

	Avg_actual_ratings	Avg_predicted_ratings
prod_id		
0594451647	0.003247	0.003360
0594481813	0.001948	0.005729
0970407998	0.003247	0.008566
0972683275	0.012338	0.035330
1400501466	0.012987	0.006966

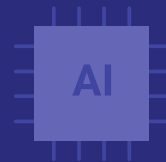
Le modèle SVD semble être une approche robuste et précise pour la recommandation de produits. Des évaluations complémentaires pourraient fournir des informations plus détaillées sur la manière dont le modèle répond aux besoins spécifiques des utilisateurs et sur d'éventuelles améliorations à apporter.



06

Conclusion

- Conclusion et perspectives



Conclusion

En conclusion, notre projet de système de recommandation de produits a démontré son efficacité à surmonter des défis tels que le "Problème du Démarrage à Froid". Les approches variées, allant des recommandations basées sur le classement au filtrage basé sur le contenu, ont permis d'offrir des recommandations personnalisées et pertinentes. Ces résultats témoignent de la robustesse de notre méthodologie face à la diversité des utilisateurs et des produits, renforçant ainsi son applicabilité dans le domaine du commerce électronique.

En perspective, envisager le déploiement de notre système de recommandation sous forme d'API web, offrant ainsi une accessibilité optimale et une intégration aisée dans les plateformes en ligne