

Moteur de recommandations

- ▶ Idée : effectuer des suggestions pertinentes pour un utilisateur
- ▶ Données disponibles :
 - ▶ Historiques des utilisateurs
 - ▶ Informations sur les suggestions disponibles
 - ▶ Liens entre les utilisateurs
- ▶ Challenges :
 - ▶ Peu de données sur chaque utilisateur
 - ▶ Beaucoup de suggestions différentes possibles
- ▶ Exemples : Amazon, Netflix, publicité ciblée

Recommended Based on Your Browsing History



Complétion de matrice

- ▶ Pour chaque contenu consommé, un utilisateur émet un avis
 - ▶ Explicite : note sur 5, j'aime
 - ▶ Implicite : durée de consultation d'une page, d'une vidéo

	C_1	C_2	C_3	C_4
U_1	4	5	—	—
U_2	3	—	3	—
U_3	—	—	—	1

Complétion de matrice

- ▶ Pour chaque contenu consommé, un utilisateur émet un avis
 - ▶ Explicite : note sur 5, j'aime
 - ▶ Implicite : durée de consultation d'une page, d'une vidéo
- ▶ Prédiction des avis qu'émettrait un utilisateur pour les contenus non consommés

	C_1	C_2	C_3	C_4
U_1	4	5	3	5
U_2	3	3	3	1
U_3	3	1	2	1

Complétion de matrice

- ▶ Pour chaque contenu consommé, un utilisateur émet un avis
 - ▶ Explicite : note sur 5, j'aime
 - ▶ Implicite : durée de consultation d'une page, d'une vidéo
- ▶ Prédiction des avis qu'émettrait un utilisateur pour les contenus non consommés
- ▶ Suggestion des contenus aux avis prédits les plus favorables

	C_1	C_2	C_3	C_4
U_1	4	5	3	5
U_2	3	3	3	1
U_3	3	1	2	1

Méthodes basées sur le contenu

- ▶ Principe : suggérer des objets proches de ceux appréciés dans le passé

Méthodes basées sur le contenu

- ▶ Principe : suggérer des objets proches de ceux appréciés dans le passé
- ▶ Méthode :
 - ▶ Caractériser les contenus pour définir une mesure de similarité
 - ▶ $s(\text{Harry Potter 1}, \text{Harry Potter 2}) = 0.9$
 - ▶ $s(\text{Harry Potter 1}, \text{Annie Hall}) = 0.1$
 - ▶ Utiliser l'historique des contenus de l'utilisateur pour prédire les avis

Méthodes basées sur le contenu

- ▶ Principe : suggérer des objets proches de ceux appréciés dans le passé
- ▶ Méthode :
 - ▶ Caractériser les contenus pour définir une mesure de similarité
 - ▶ $s(\text{Harry Potter 1}, \text{Harry Potter 2}) = 0.9$
 - ▶ $s(\text{Harry Potter 1}, \text{Annie Hall}) = 0.1$
 - ▶ Utiliser l'historique des contenus de l'utilisateur pour prédire les avis
- ▶ Avantages :
 - ▶ Peu coûteux
 - ▶ Nouveau contenu peut être recommandé

Méthodes basées sur le contenu

- ▶ Principe : suggérer des objets proches de ceux appréciés dans le passé
- ▶ Méthode :
 - ▶ Caractériser les contenus pour définir une mesure de similarité
 - ▶ $s(\text{Harry Potter 1}, \text{Harry Potter 2}) = 0.9$
 - ▶ $s(\text{Harry Potter 1}, \text{Annie Hall}) = 0.1$
 - ▶ Utiliser l'historique des contenus de l'utilisateur pour prédire les avis
- ▶ Avantages :
 - ▶ Peu coûteux
 - ▶ Nouveau contenu peut être recommandé
- ▶ Inconvénients :
 - ▶ Quelle similarité ?
 - ▶ N'utilise que les données de l'utilisateur courant
 - ▶ Impossible d'effectuer une recommandation pour un nouvel utilisateur

Méthodes collaboratives historiques

- ▶ Principe : suggérer des contenus appréciés par les utilisateurs aux goûts similaires

Méthodes collaboratives historiques

- ▶ Principe : suggérer des contenus appréciés par les utilisateurs aux goûts similaires
- ▶ Méthode :
 - ▶ Déterminer les similarités entre utilisateurs à partir des avis sur des contenus communs
 - ▶ Effectuer une prédiction grâce aux avis déjà émis par des utilisateurs similaires

Méthodes collaboratives historiques

- ▶ Principe : suggérer des contenus appréciés par les utilisateurs aux goûts similaires
- ▶ Méthode :
 - ▶ Déterminer les similarités entre utilisateurs à partir des avis sur des contenus communs
 - ▶ Effectuer une prédiction grâce aux avis déjà émis par des utilisateurs similaires
- ▶ Avantages :
 - ▶ Utilise l'ensemble des données à disposition
 - ▶ Ne nécessite pas de compréhension des contenus

Méthodes collaboratives historiques

- ▶ Principe : suggérer des contenus appréciés par les utilisateurs aux goûts similaires
- ▶ Méthode :
 - ▶ Déterminer les similarités entre utilisateurs à partir des avis sur des contenus communs
 - ▶ Effectuer une prédiction grâce aux avis déjà émis par des utilisateurs similaires
- ▶ Avantages :
 - ▶ Utilise l'ensemble des données à disposition
 - ▶ Ne nécessite pas de compréhension des contenus
- ▶ Inconvénients :
 - ▶ Impossible de recommander un nouveau contenu
 - ▶ Impossible d'effectuer une recommandation pour un nouvel utilisateur

Méthodes collaboratives avec modèle

- ▶ Principe : apprendre un modèle pour prédire les avis

Méthodes collaboratives avec modèle

- ▶ Principe : apprendre un modèle pour prédire les avis
- ▶ Méthode :
 - ▶ Définir un modèle prédictif
 - ▶ ex : $avis(u, c) = m + f(u) + g(c)$
 - ▶ Utiliser l'ensemble des données pour apprendre les paramètres du modèle et prédire les avis manquants

Méthodes collaboratives avec modèle

- ▶ Principe : apprendre un modèle pour prédire les avis
- ▶ Méthode :
 - ▶ Définir un modèle prédictif
 - ▶ ex : $avis(u, c) = m + f(u) + g(c)$
 - ▶ Utiliser l'ensemble des données pour apprendre les paramètres du modèle et prédire les avis manquants
- ▶ Avantages :
 - ▶ Utilise l'ensemble des données à disposition
 - ▶ Bonne qualité de prédiction lorsque modèle adapté

Méthodes collaboratives avec modèle

- ▶ Principe : apprendre un modèle pour prédire les avis
- ▶ Méthode :
 - ▶ Définir un modèle prédictif
 - ▶ ex : $avis(u, c) = m + f(u) + g(c)$
 - ▶ Utiliser l'ensemble des données pour apprendre les paramètres du modèle et prédire les avis manquants
- ▶ Avantages :
 - ▶ Utilise l'ensemble des données à disposition
 - ▶ Bonne qualité de prédiction lorsque modèle adapté
- ▶ Inconvénients :
 - ▶ Souvent coûteux
 - ▶ Impossible de recommander un nouveau contenu
 - ▶ Impossible d'effectuer une recommandation pour un nouvel utilisateur

Complétion de matrice

- ▶ Méthodes variées
- ▶ Adaptées à un grand nombre de problèmes
- ▶ Plusieurs inconvénients communs :
 - ▶ Pas de recommandation pour un nouvel utilisateur
 - ▶ Ne propose que du contenu similaire
- ▶ Regarder les approches par bandits (exploration/exploitation)