

RECOMMANDATION DE CONTENU

Antonin Garret



My Content

Sommaire

- I. Présentation du problème
- II. Modèle de recommandation :
 - a. Collaborative Filtering
 - b. Content-based Filtering
- III. Application Test
- IV. Architecture cible

I. Présentation du problème

Développement d'une application de recommandation

Basé sur les données de **Globo.com** :

- Base de données d'articles de presse
- 364047 articles, 322897 utilisateurs
- Données sur les interactions des utilisateurs avec les articles

Prévoir une architecture pour l'application :

- API pour accéder aux recommandations
- Ajout d'un nouvel utilisateur
- Ajout d'un nouvel article

Séparation des cliques selon les dates

- Entrainement : 80% des cliques les plus anciens
- Validation : 20% des cliques les plus récents

Test des modèles :

- Recommandation de 5 articles par utilisateur
- Précision de la recommandation :
 - Nombre articles recommandés lus par l'utilisateur dans le jeu de validation / 5

Sélection des utilisateurs :

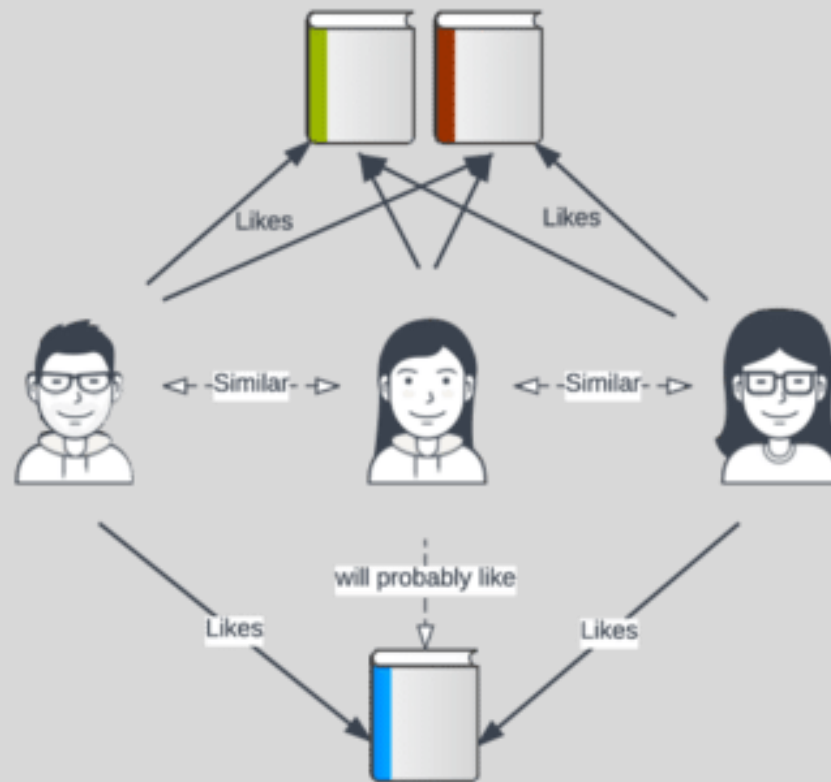
- Ayant lu au moins 5 articles dans le jeu de validation
- Ayant lu au moins 1 articles dans le jeu d'entraînement
- 20 497 utilisateurs

II. Modèle de recommandation

Méthodologie

II. Modèle de recommandation

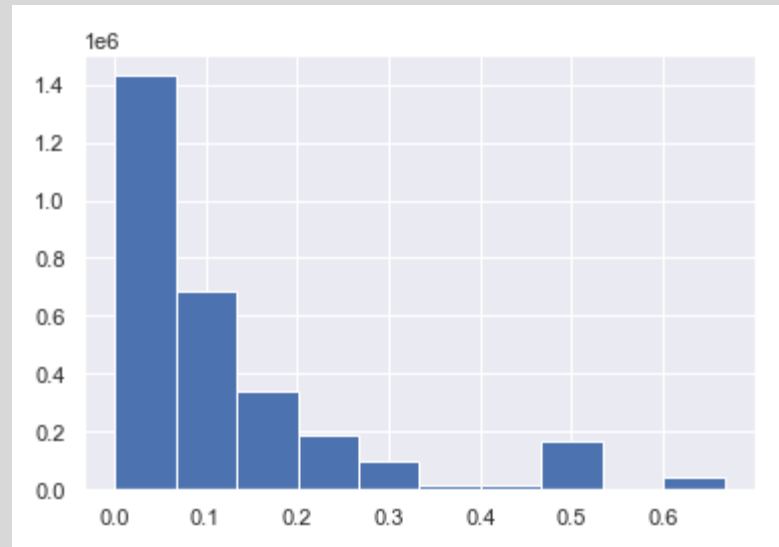
Collaborative Filtering



II. Modèle de recommandation

Collaborative Filtering

- Métrique pour évaluer l'attrait d'un utilisateur pour un article :
 - *art_rating* : Cliques de l'utilisateur sur l'articles / cliques totaux de l'utilisateur
 - *cat_rating* : Cliques de l'utilisateur sur la catégorie de l'article / cliques totaux de l'utilisateur
 - *hmean_rating* : Moyenne harmonique de ces deux métriques



Distribution de la métrique *hmean_rating*

II. Modèle de recommandation

Collaborative Filtering

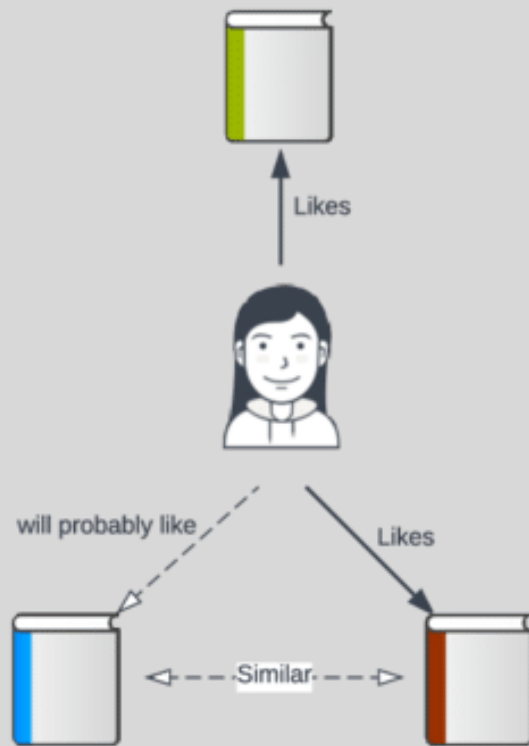
Articles recommandés : 5 articles non lus ayant le score prédit le plus haut

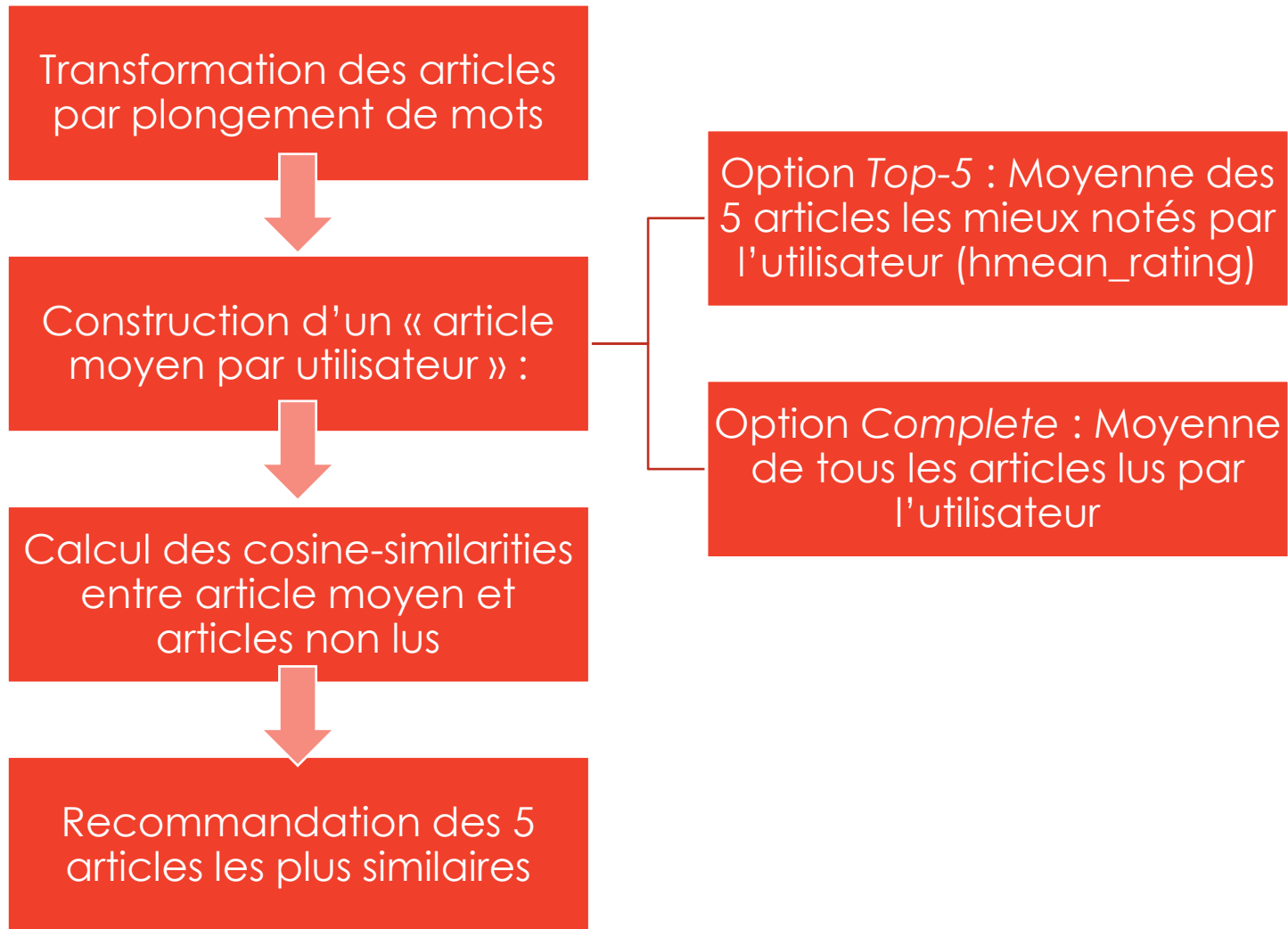
Modèle	SVD	KNN-means
RMSE	0.0499	0.0164
Précision moyenne	0.000196	0.000196

Recommandations identiques entre SVD et KNN-means

II. Modèle de recommandation

Content-based Filtering





02/12/2022

II. Modèle de recommandation

Content-Based Filtering

Résultats :

Approche	Précision moyenne
Top-5	0.00711
Complete	0.00421

III. Application Test

Architecture serverless : Azure fonction



- Pas de gestion de l'infrastructure
- Mise a l'échelle automatique
- Paiement à l'utilisation

Activation par Trigger :

- Recommandation :
Trigger par appel à une url
- Route :
`{url}/api/user/{user_id}`

III. Application Test API



02/12/2022

III. APPLICATION TEST

Démonstration

IV. Architecture cible

