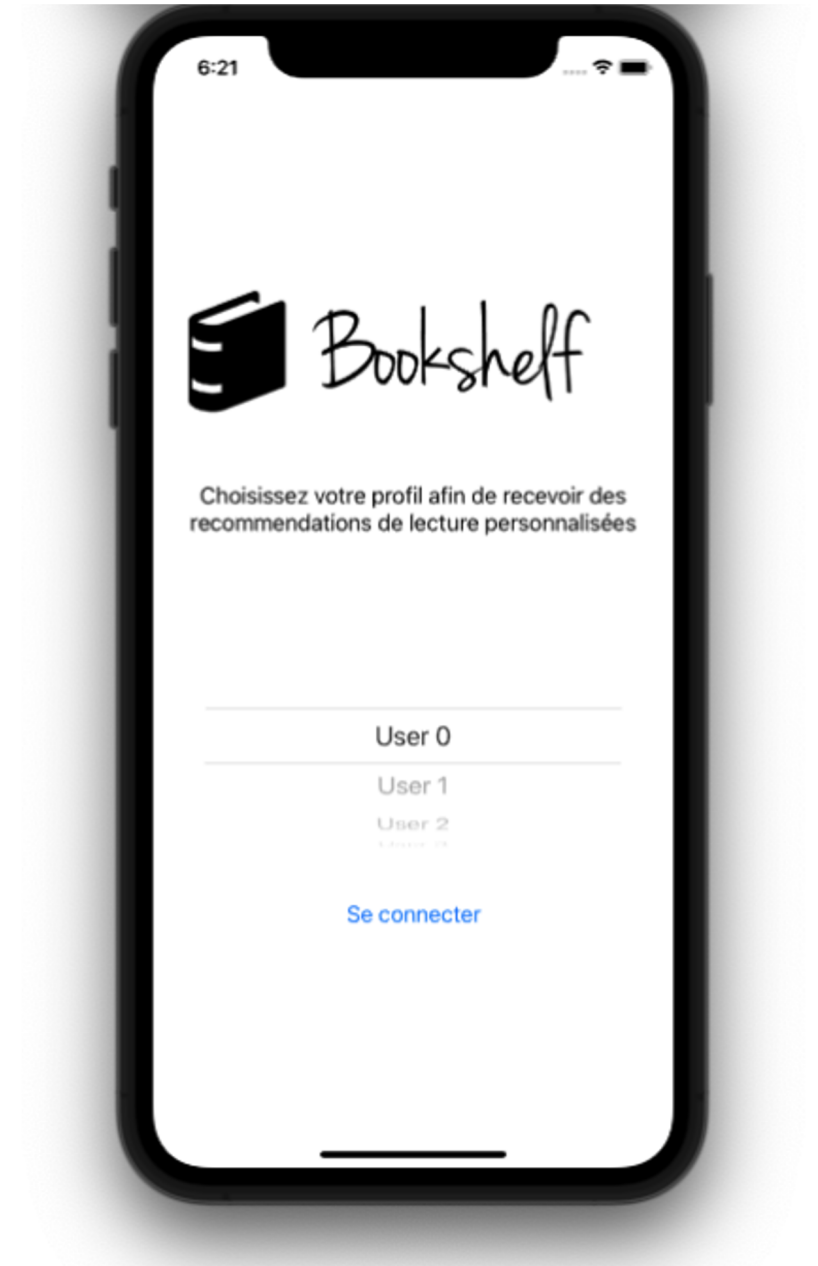




Systeme de recommandation de contenu



Description fonctionnelle

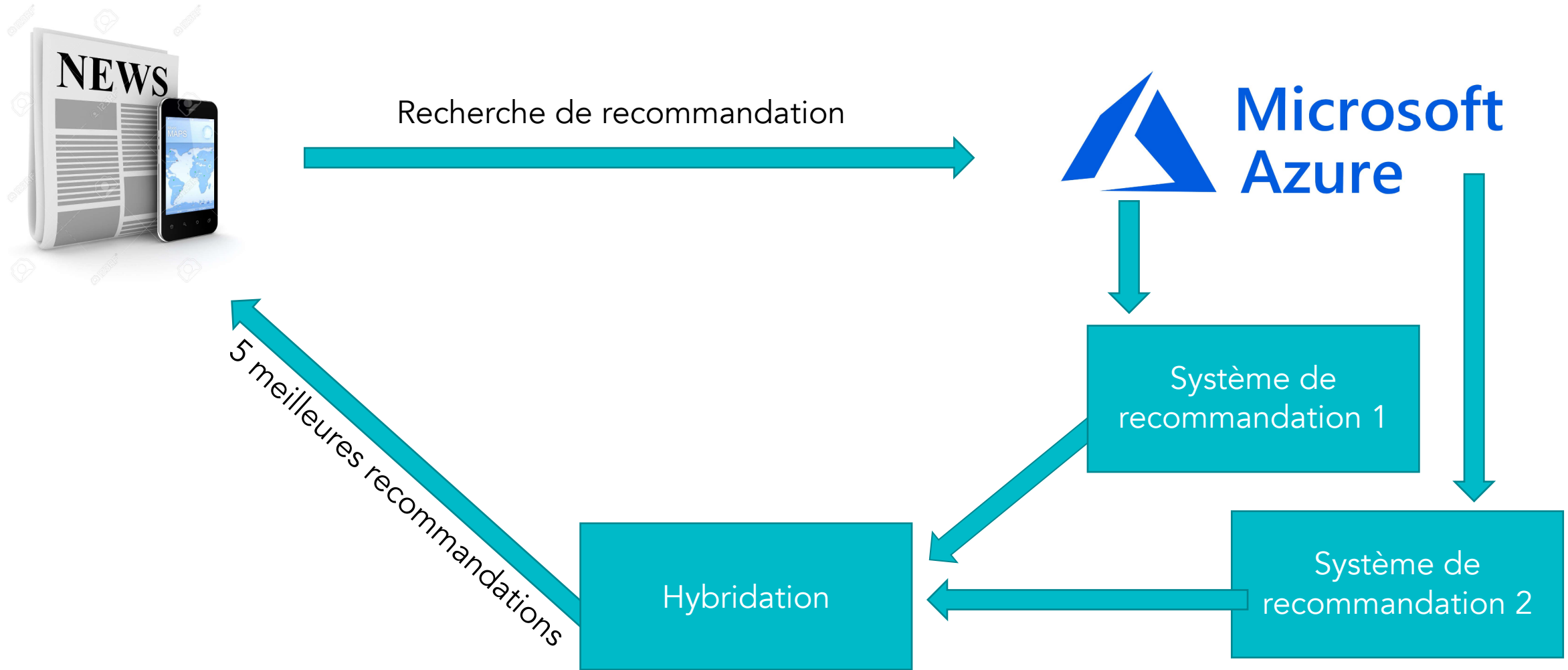
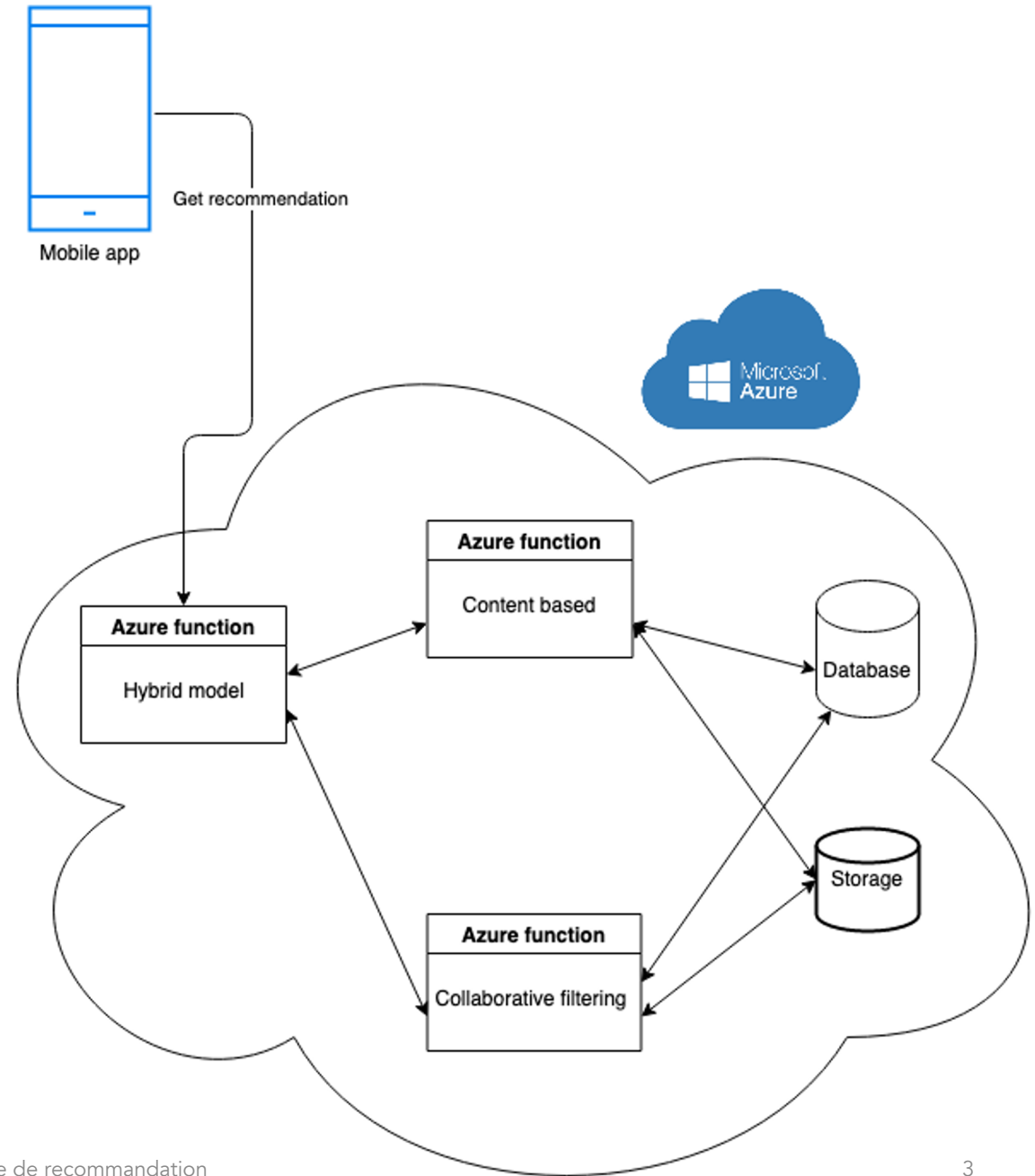
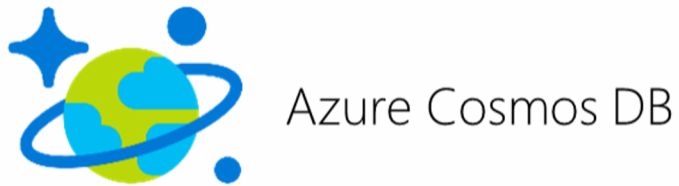
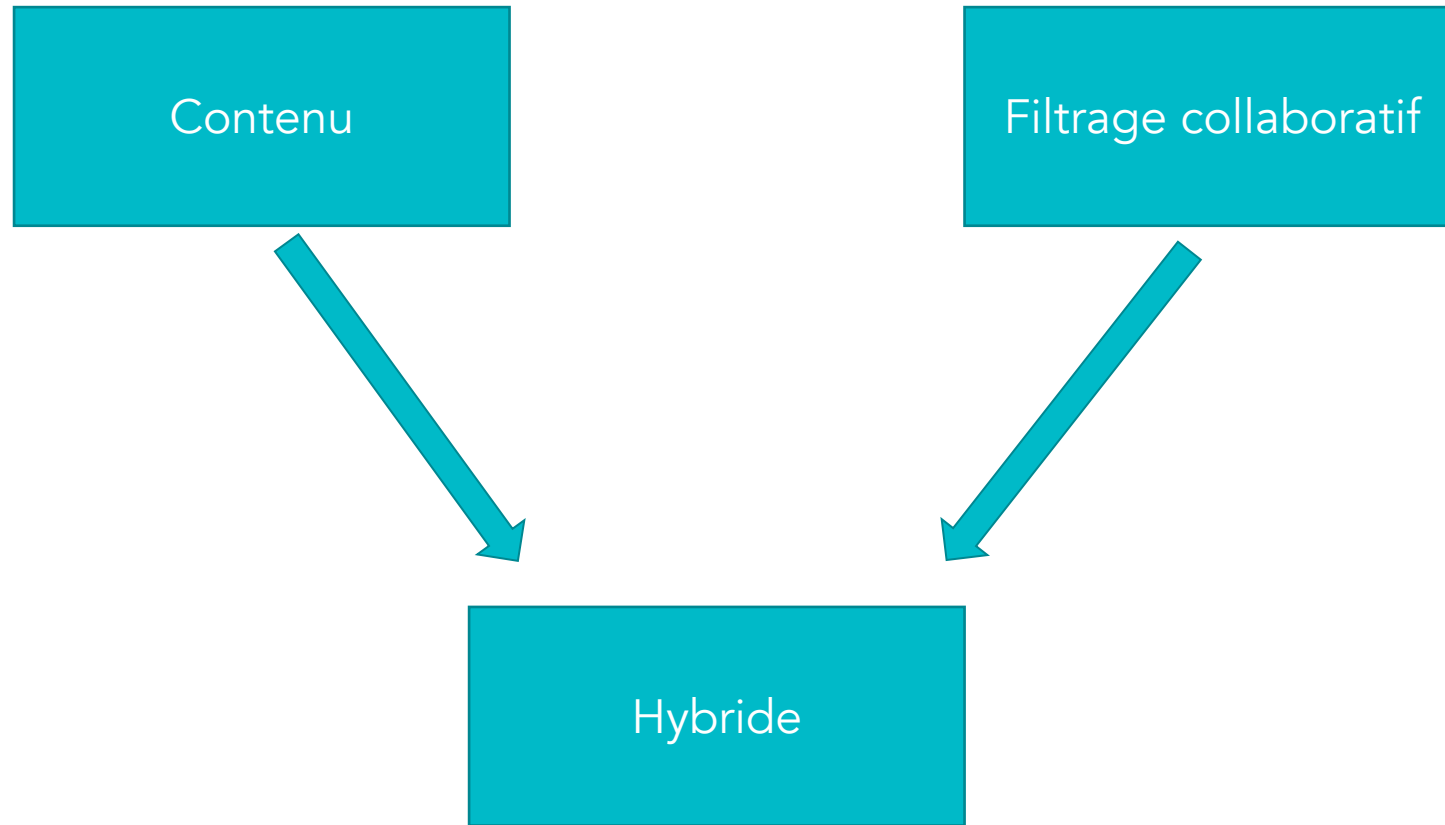


Schéma de l'architecture



Système de recommandation



Filtrage basé sur le contenu

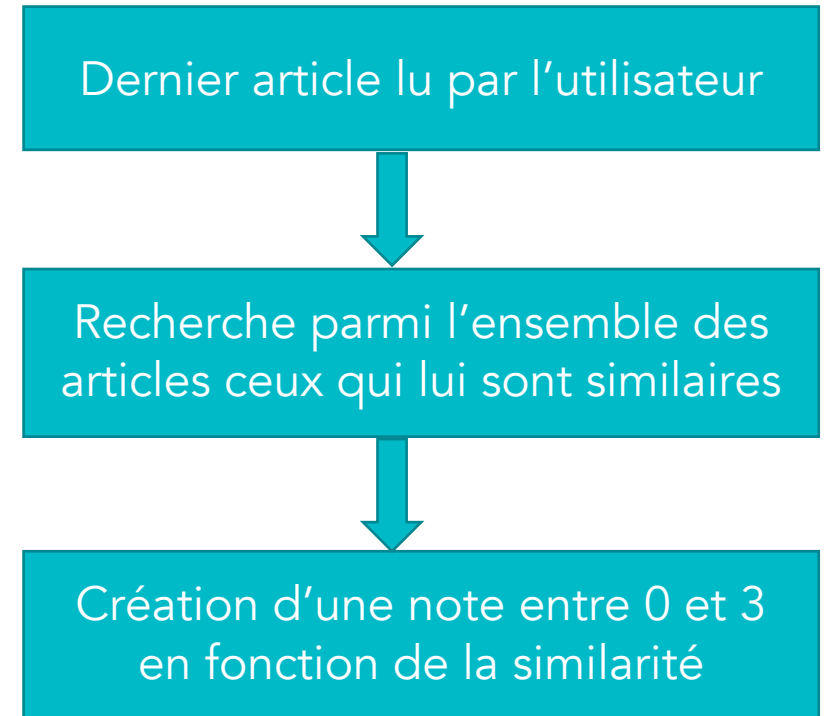
Système basé sur le profil de l'utilisateur

Avantages

- Pas besoin de données sur les autres utilisateurs.
- Possibilité de recommander des produits qui ne sont pas populaires ou nouveaux.
- Possibilité de recommander aux utilisateurs ayant un goût unique ou rare.

Inconvénients

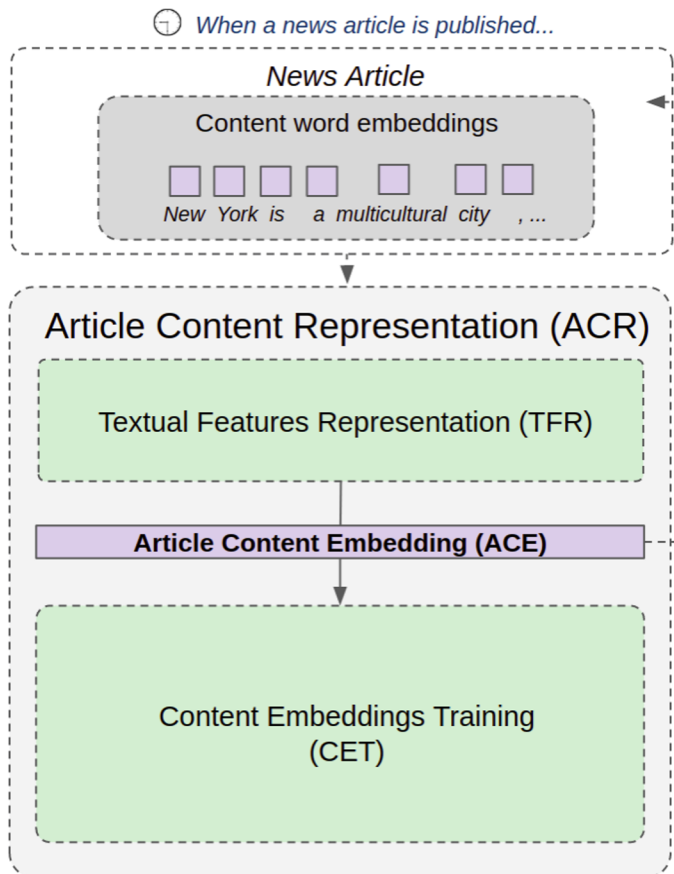
- Trouver les bons "features" n'est pas toujours facile.
- Comment créer un profil pour les nouveaux utilisateurs ?



Filtrage basé sur le contenu

Création du système de similarité

CHAMELEON - A Deep Learning Meta-Architecture for News Recommender Systems



Le module **ACR** est chargé d'extraire des caractéristiques du texte et des métadonnées des articles de presse et d'apprendre des représentations distribuées (embeddings) pour chaque contexte d'article de presse.

Le sous-module **TFR** est chargé d'apprendre les caractéristiques pertinentes directement à partir du contenu textuel de l'article, et peut être instancié en utilisant des CNN et des RNN, par exemple.

Le sous-module **CET** est responsable de l'apprentissage des Article Content Embeddings (ACE) pour une tâche secondaire. Il peut être instancié comme un modèle d'apprentissage supervisé ou non supervisé.

Filtrage collaboratif

Système basé sur les intérêts d'un groupe d'utilisateurs

Avantages

- Pas d'information sur les produits et leurs spécialités.
- Simplicité de corrélation entre les produits.

Recherche dans la matrice les articles les plus pertinents pour cet utilisateur par rapport aux autres utilisateurs

Inconvénients

- Un nouvel utilisateur commence sans avoir une recommandation
- Dans le cas d'un large nombre de produits, il devient difficile de "matcher" les utilisateurs qui ont les mêmes préférences.

Filtrage collaboratif

Création du système de similarité

Surprise est un module pour construire et analyser des systèmes de recommandation qui traitent des données de notation explicites.

Création d'un système de notation

Nombre de clics de l'utilisateur divisé par le nombre d'articles vu pendant une session de l'utilisateur

Factorisation de matrice

Matrice où chaque ligne représente un utilisateur et chaque colonne un article. Les éléments de cette matrice sont les notes attribuées aux articles par les utilisateurs.

Hybridation

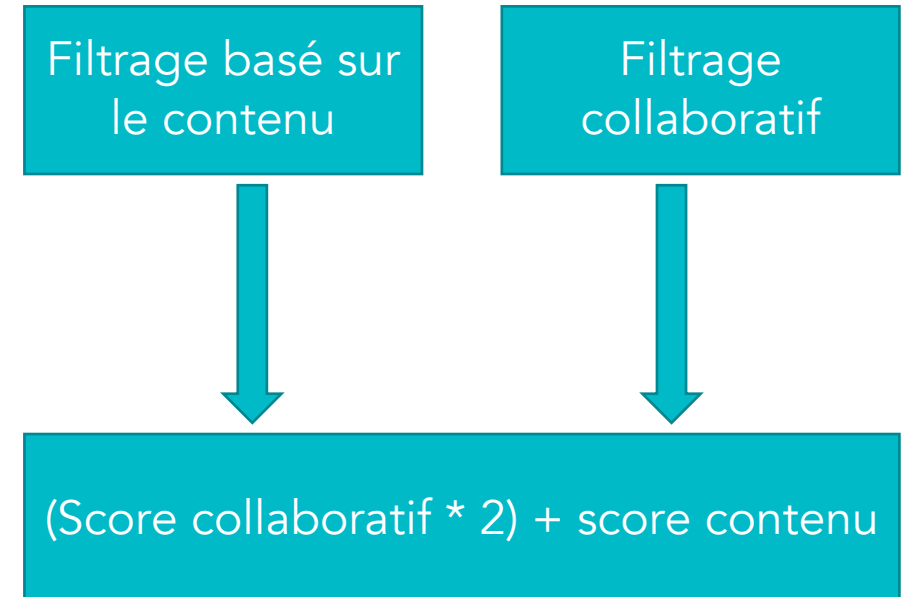
Système basé sur le filtrage sur le contenu et le filtrage collaboratif

Avantages

- Meilleurs des deux méthodes

Inconvénients

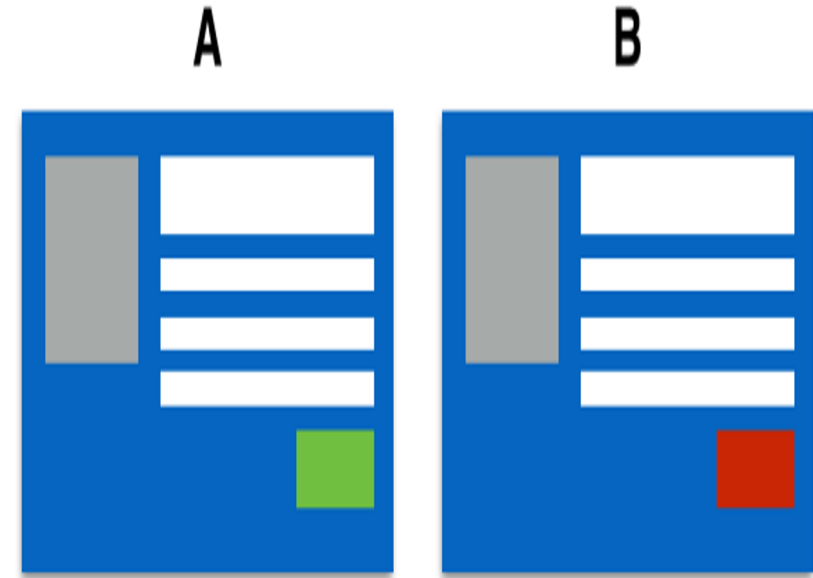
- Temps de traitement allongé



Tests



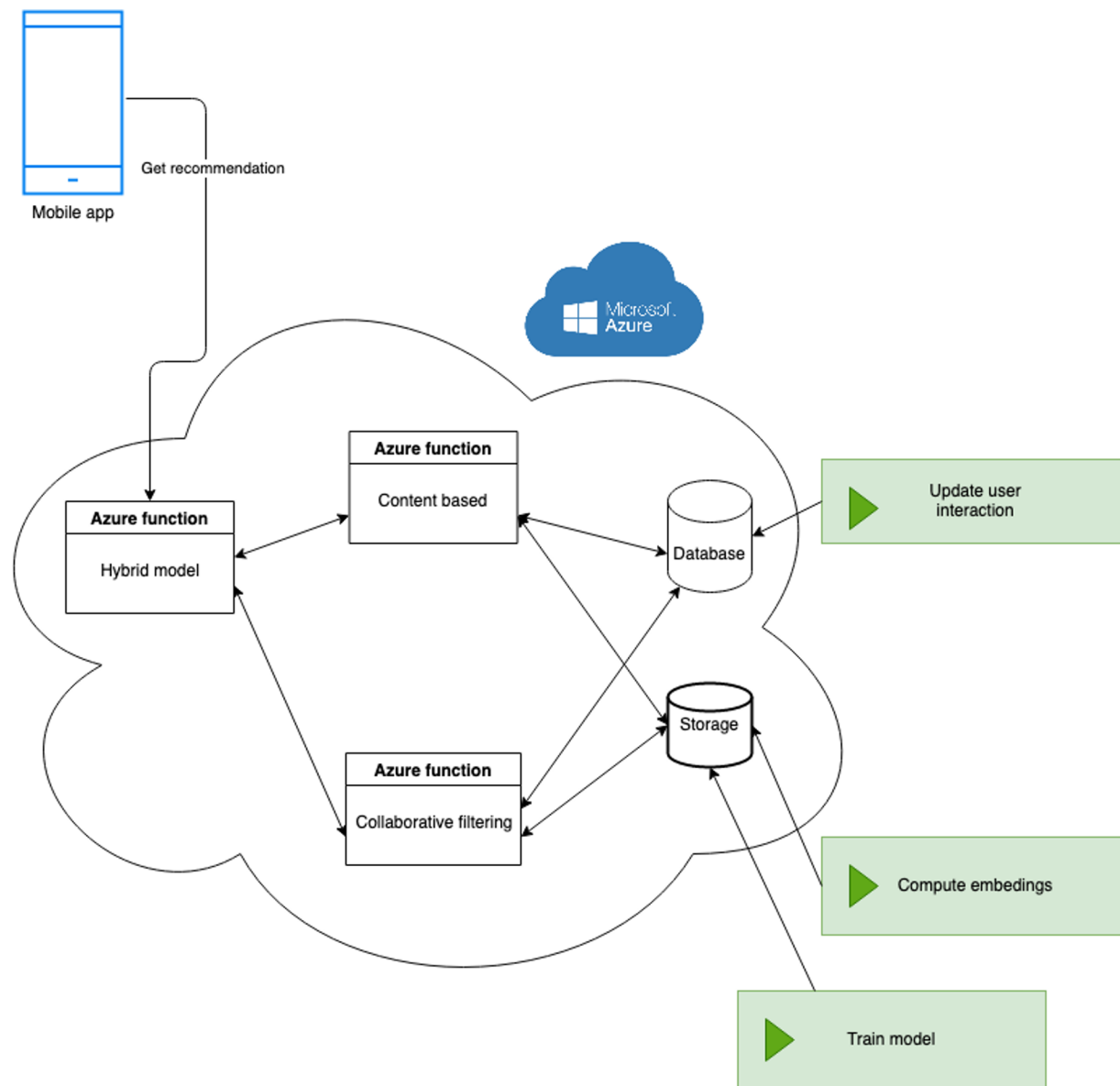
Mise en place de tests sur les données passées



Mise en place d'un système A/B test

Architecture complète

- Système de recommandation
- Alimentation des données



Sources

Application mobile

<https://github.com/OC-Ingenieur-IA/bookshelf?organization=OC-Ingenieur-IA&organization=OC-Ingenieur-IA>

Système de recommandations

<https://github.com/PierreSylvain/MyContentFunctions>

Exploration des données

<https://colab.research.google.com/drive/16ka6fPjp1NgSHxFfcGn3sKQT-g18clKY?usp=sharing>

Recommandations basées sur le contenu

<https://colab.research.google.com/drive/1Dj6UaIT6mCeLJJMnas4tz8i7sCPEwOUq?usp=sharing>

Recommandations basées sur le filtrage collaboratif

<https://colab.research.google.com/drive/1yhFVqrO7QdtpmlAnD79zJQJgMeHiP919?usp=sharing>

Recommandations basées sur un système hybride

https://colab.research.google.com/drive/1H4dd6iBFWl_yhOLADnASUg-jrEySDxkY?usp=sharing

Questions

