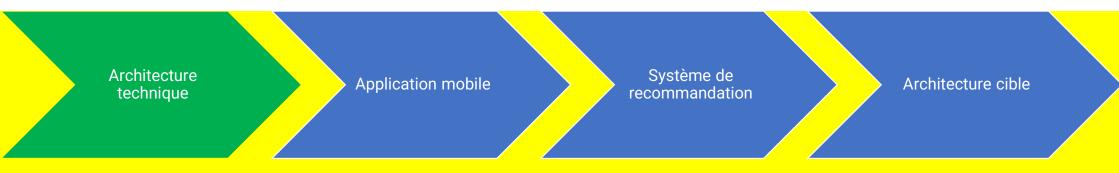


SYNTHESE

PROJET 9 : Réalisez une application mobile de recommandation de contenu

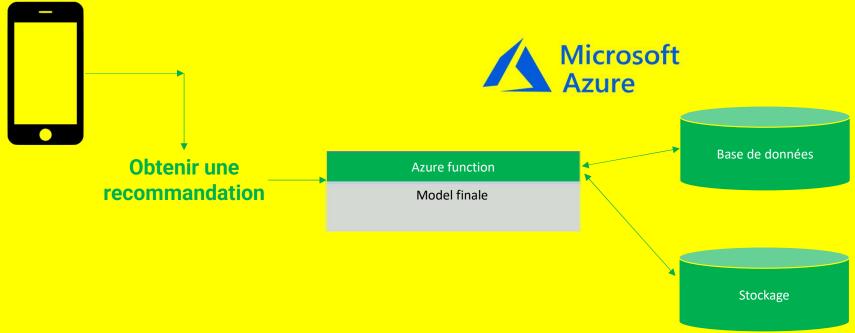
Une application mobile de recommandation de contenu basée sur des articles chez "OPENCLASSROOM"

Jaoid KRAIRI (Mai 2022)



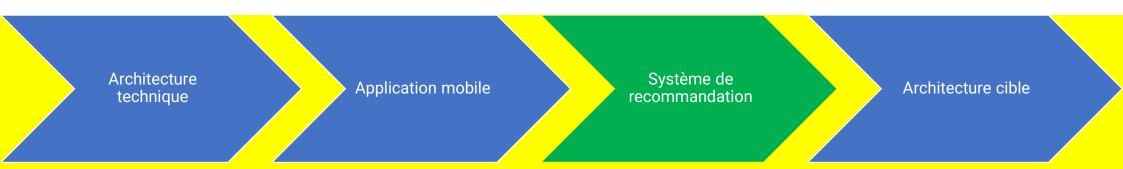
Le lien entre l'application et le système de recommandation utilise une architecture Serverless.

Plus particulièrement le service Azure Functions qui permet de mettre en place rapidement ce type d'architecture dans le cloud.



L'application mobile de recommandation de contenu appelle un service web de type REST qui est une application Serverless disponible sur le cloud Microsoft Azure. L'application mobile transmet l'identifiant de l'utilisateur à la fonction de recommandation de contenu et cette fonction renvoie les identifiants des 5 articles recommandés pour cet utilisateur.



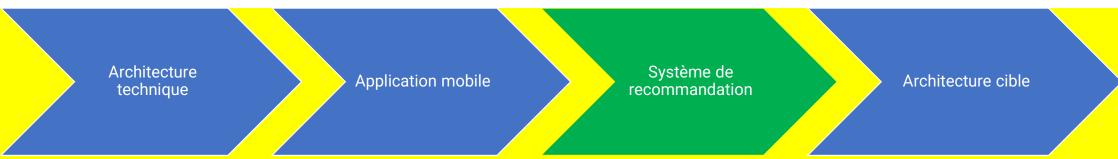


La construction de mes systèmes de recommandations a été réalisée à l'aide de scikit-surprise (bibliothèque surprise). Les systèmes de recommandation sont l'une des applications les plus couramment utilisées et facilement compréhensibles de la science des données.

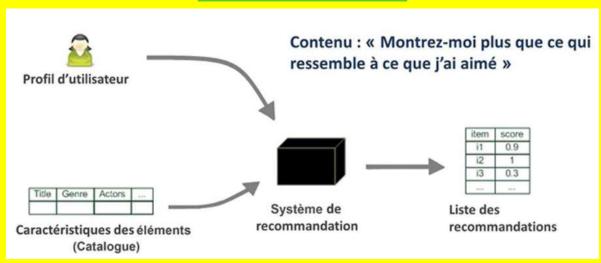
Beaucoup de travail a été fait sur ce sujet, l'intérêt et la demande dans ce domaine restent très élevés en raison de la croissance rapide d'Internet et du problème de surcharge d'informations.

Il est devenu nécessaire pour les entreprises en ligne d'aider les utilisateurs à faire face à la surcharge d'informations et de leur fournir des recommandations de contenu dans mon cas personnalisés.

Deux des façons les plus populaires d'aborder les systèmes de recommandation sont les recommandations basées sur le contenu et le filtrage collaboratif.

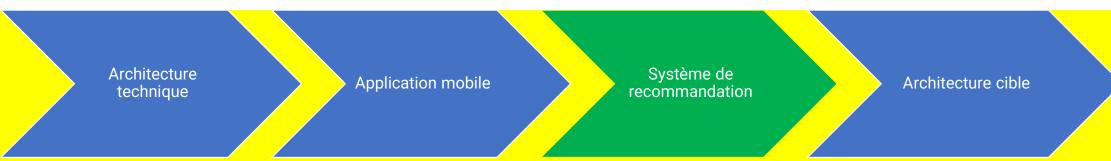


Basées sur le contenu

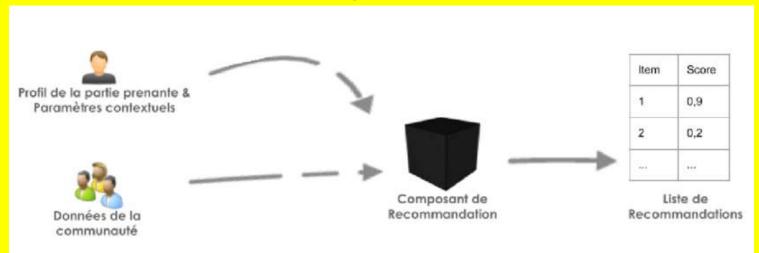


Pour les recommandations basées sur le contenu, la tâche consiste à déterminer quels éléments du catalogue coïncident le mieux avec les préférences de l'utilisateur. Une telle approche ne requiert pas une grande communauté d'utilisateurs ou un gros historique d'utilisation du système. Elle va prendre le dernier article lu par l'utilisateur et recherché parmi tous les articles ceux qui sont le plus proche de cet article en terme de caractéristiques.

Néanmoins, le manque d'information ne me permet pas d'utiliser cette méthode de manière efficace.

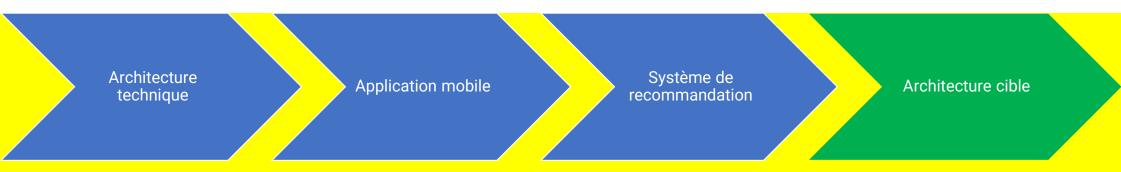


Le filtrage collaboratif



Le filtrage collaboratif est un algorithme de recommandation, c'est-à-dire un algorithme qui consiste à prédire les articles dans mon cas que des utilisateurs apprécieront dans le futur. Il repose sur l'exploitation de « la sagesse de la foule ». L'idée est que pour prédire ce qu'une personne va aimer dans le futur, et qu'il est donc pertinent de lui recommander, l'algorithme n'utilise pas uniquement l'historique de cet utilisateur, mais toutes les informations existantes dans le système concernant les autres utilisateurs. Pour faire simple, ces algorithmes permettent de détecter les utilisateurs qui ont des goûts similaires pour exploiter ces informations à des fins de recommandations. Dans mon cas, j'ai utiliser et évaluer 2 algorithmes, l'un était le SVD qui est équivalent à la factorisation matricielle probabiliste et l'autre était le KNNWithMeans qui prend en compte les notes moyennes de chaque utilisateur.

Au final, j'ai garder le KNNWithMeans car il apporte de meilleur performance.



Le système de recommandation nécessite d'être alimenté par les nouvelles interactions des utilisateurs et la publication de nouveaux articles. Les nouvelles interactions des utilisateurs sont stockées dans une base de données hébergée sur le cloud grâce à l'API Azure Cosmo DB pour MongoDB.

Lorsqu'un nouvel article est publié une application se chargera de créer les intégrations et stockera le résultat dans un espace de stockage Azure.

À la fin de chacun des deux processus (interaction et ajout d'un article), l'algorithme Surprise sera lancé pour mettre à jour les données.

Les traitements pour la création des intégrations et l'algorithme Surprise sont couteux en temps de traitement et de ressource.

Il est possible d'envisager un traitement en lot pour minimiser les coûts.

