

# Algorithmes de recommandation et libre arbitre

A l'heure du numérique, l'explosion de la quantité d'information accessible a conduit à la nécessité d'implémenter des algorithmes de recommandation qui filtrent le contenu proposé à chacun. Cependant, ces algorithmes soulèvent des enjeux impliquant individu comme communauté dans son rapport au numérique : manque de contraste lié à une information ciblée et création de comportements culturels grégaires et normalisés.

## I Fonctionnement des algorithmes de recommandation

### I.1 Collection de données

Comment les algorithmes de recommandation filtrent et collectent nos données ?

Avant de pouvoir nous proposer une liste de recommandations, un algorithme cherche à collecter un certain nombre d'informations pour caractériser et donner un profil à l'utilisateur.

On peut collecter des données explicitement : ce type de collecte va directement recueillir les informations à l'utilisateur, par exemple en prenant en compte ce qu'il a pu "liker" sur certains réseaux, s'il a partagé tel ou tel contenu... Ce type de collecte, contrairement au second cité ci-dessous, a l'avantage d'identifier clairement l'utilisateur, de pouvoir faire la différence entre plusieurs personnes qui utiliseraient des applications à partir du même téléphone ou ordinateur.

La collecte de données implicites : Les données de l'utilisateur vont être collectées sans même que celui ne s'en rende compte. Cette collecte se base sur l'analyse des positions GPS pour certaines applications comme Snapchat, Instagram ou encore Tinder, sur les listes de musique écoutées par l'utilisateur pour des plateformes musicales comme Deezer ou Spotify, ou encore sur l'analyse de la fréquence à laquelle un utilisateur se connecte sur une page. Bien que ces données soient facilement collectées, il n'en est pas moins facile d'attribuer ces données à un utilisateur comme cité plus haut.

## I.2 Stratégies de recommandation

Les données une fois collectées, il vient de les exploiter afin de recommander le meilleur contenu. On distingue trois méthodes de recommandation, se démarquant par les catégories de données utilisées principalement :

- Basée sur le contenu : ici, on s'intéresse essentiellement au contenu consommé par l'utilisateur. Il s'agit de proposer à l'utilisateur des produits similaires à ceux précédemment consommés. Par exemple, le temps passé sur une vidéo Youtube, le nombre d'écoutes d'une musique sur Spotify, ou encore un commentaire laissé sur un produit acheté sur Amazon sont des informations importantes pour un tel algorithme. Une telle stratégie a comme avantage le fait qu'elle tient compte uniquement de l'expérience de l'utilisateur seul. En revanche, il est difficile de recommander une nouvelle personne, n'ayant pas encore beaucoup, voire pas du tout consommé. On appelle cela "le cold-start".
  - Basée sur le filtrage collaboratif : ce type d'algorithme s'intéresse à regrouper des utilisateurs similaires. Il est possible de rapprocher des utilisateurs par leur consommation : deux utilisateurs seront proches s'ils consomment les mêmes produits. On peut également grouper deux utilisateurs en effectuant des similitudes entre leur informations personnelles, deux personnes ayant des situations géographiques proches par exemple, ou tout simplement par des amis ou abonnements communs, sur des plateformes comme Facebook et Twitter. Cela pousse certaines structures ne se présentant pas réellement comme un réseau social à se développer comme tel : il est par exemple possible de s'abonner au profil des utilisateurs sur Spotify. Ces rapprochements faits, l'algorithme va alors proposer des contenus consommés par des personnes similaires à l'utilisateur. Cela peut se résumer par la phrase "Un tel a aimé ceci, vous aimerez sûrement cela...".
  - Recommandation de produits populaires : cette méthode consiste à recommander les contenus les plus populaires, la définition de ce terme changeant en fonction de chaque algorithme. Par exemple, pour l'algorithme PageRank de Google, lors d'une recherche ciblée, il s'agit de proposer les pages webs les plus référencées par d'autre. Concrètement, la page se voit attribuer un score, en fonction du nombre de fois qu'un utilisateur se déplace de manière aléatoire sur le web rencontrerait la page en question. On peut supposer qu'un produit populaire aura de bonnes chances de plaire, mais cette méthode est moins efficace pour recommander un utilisateur ayant des préférences plus originales.
- + sur Page Rank

## 2 Conséquences

Les algorithmes de recommandation utilisant un modèle d'utilisateur déterminé sur les comportements précédant filtrent l'information en confortant l'avis de l'utilisateur : c'est la bulle filtrante. Ce procédé émancipe et accentue un phénomène bien connu du monde de la communication, l'effet chambre d'écho. Le succès de cette méthode de recommandation utilise le biais de confirmation, biais cognitif poussant le consommateur de contenu à consulter toujours plus d'information en accord avec ces idéaux.

Finale la théorie de la bulle filtrante, bien qu'un phénomène social difficilement mesurable et contesté, possède ses partisans à tout niveaux sociaux et de professionnalisation. De nombreux témoignages du consommateur amateur d'informaticaux aux ingénieurs et hommes politiques soutiennent l'existence de cette enjeu que soulève la connexions de plus en plus intenses entre les hommes qui finit par les séparer (+ EX BULLE FILTRANTE)

Un des problèmes posés par le processus d'automatisation et d'algorithmisation de notre rapport au numérique est celui de l'individualisation : en nous donnant la voie vers un monde non seulement plus vaste mais aussi plus facile d'accès et dont la navigation est plus aisée, est-il plus facile de s'émanciper avec le numérique ? Effectivement, l'internet nous donne accès au monde mais aussi donne au monde l'accès à nous : par le biais des notifications par exemple, il est bien plus facile d'attirer l'attention d'un individu et c'est une conséquence de la numérisation du monde que les entreprises ont transformé en atout, comme en atteste cette citation fameuse du directeur de TF1, Patrick Lelay : « Soyons réalistes : le business de TF1, c'est de vendre du temps de cerveau disponible à Coca Cola ».

Selon Bernard Stiegler, philosophe français, on assiste avec l'algorithmisation du numérique à une forme de « marketing de la libido » (selon lui, ce terme désigne le désir de l'individu ainsi que ses pulsions) : les entreprises cherchent par tous les moyens à capter notre attention (par exemple par la télévision, par les publicités) : on crée le désir de consommer en présentant le produit comme le moyen de satisfaire ce dernier. Et ce phénomène a été facilité par l'arrivée du numérique : la consommation moderne repose sur des principes de recommandation par des pairs. Ces algorithmes conduisent à l'émergence d'un désir unifié puisque l'individu est redirigé vers des produits qu'il est censé aimer (ce que l'algorithme calcule et prévoit). Le désir n'est pas renouvelé, puisque l'utilisateur ne découvre pas véritablement de produit nouveau : en se laissant guider par la lecture automatique de Youtube par exemple, les vidéos qu'il consomme sont sélectionnées pour être fortement susceptibles de lui plaire : au lieu de découvrir de nouvelles choses, l'utilisateur est enfermé dans une culture qui peut être aussi peu diversifiée qu'elle en devient universelle (à travers la rubrique « Tendances » de Youtube, l'utilisateur est redirigé vers du contenu apprécié par tous). Ainsi, affirme Bernard Stiegler, on assiste à un véritable processus de dé-individualisation par les algorithmes...

## 3 Que faire contre ?

DuckDuckGo préserve la vie privée et ne stocke aucune information personnelle concernant les utilisateurs (adresse IP et traces numériques comme la signature de navigateur).

Le principe de ce moteur de recherche et de ne pas limiter les utilisateurs à une "bulle" d'information qui serait définie de manière personnalisée pour chaque individu (comme le font la plupart des autres moteurs de recherche) mais bien de permettre à tous d'accéder au Web tout entier. DuckDuckGo a pour fonctionnalité de proposer les mêmes résultats non filtrés à tous les utilisateurs en leur laissant le choix de régler des paramètres explicites : spécialisation géographique des résultats, masquage ou non de résultats adultes etc... Cela permet une meilleure sérénité (le fait de réaliser une découverte scientifique de façon inattendue, très souvent dans le cadre d'une recherche sur un autre sujet). + définition bulle filtrante ? Dans l'introduction ?

Exemple connu bulle filtrante (même recherche Google qui amène sur des contenus de droite/gauche)

« la représentation que se fait un individu de lui-même est construite sur une somme de contradictions et d'idéaux. Nous avons la possibilité de changer en permanence. Or ces algorithmes détruisent notre singularité en créant un double de nous-même qui n'est qu'une image statistique. » Bernard Stiegler

Quelles sont les solutions pour faire valoir son "droit à l'information" et volontairement sortir de sa bulle de filtre ?

Rencontrer des points de vue différents (opposés) au siens.

Mark Zuckerberg a minimisé l'impact et la portée de ce phénomène sur Facebook mais l'entreprise commence tout de même à travailler sur des mesures pour limiter cet impact.

Il existe aussi des extensions sur différents moteurs de recherche qui permettent de mettre en lumière les bulles de filtres.

En France, le journal Le Drenche propose pour tous les sujets deux avis opposés (cependant cela pourrait tendre à privilégier les extrêmes au détriment des points de vue nuancés).

Critique de cette pensée (« le système de sélection de Facebook ne modifie que de 1 % l'exposition aux contenus politiques de camps opposés » André Gunther dans "Et si on arrêtait avec les bulles de filtre ?", et Dominique Cardon : « la bulle, c'est nous qui la créons. Par un mécanisme typique de reproduction sociale. Le vrai filtre, c'est le choix de nos amis, plus que l'algorithme de Facebook ».

De nos jours, une solution pour pallier la concentration et au manque de diversité des contenus multimédia, bien trop souvent limités par des algorithmes de recommandation, serait de proposer des ressources nouvelles aux consommateurs, par le biais de site mettant en avant ces dernières.

Un exemple concret et parlant serait la plateforme Forgotify, qui regroupe plus de 4 millions de titres oubliés et méconnus sur Spotify. Cette solution a pour but de sortir les auditeurs d'un enfermement autour d'un style musical, de playlists proposées par les algorithmes souvent redondantes. Finalement cette plateforme permet de découvrir du contenu original, tout en permettant aux auteurs de gagner en visibilité.

Dans une optique similaire, Soundcloud est une plateforme de musique qui se veut très facilement accessible, et permet à tout type d'artiste de proposer son contenu gratuitement. Une particularité de ce média social est qu'il est très facile d'y trouver un contenu nouveau, proposé par des artistes souvent inconnus, des remix. En contrepartie les musiques et albums les plus populaires y seront souvent absents.