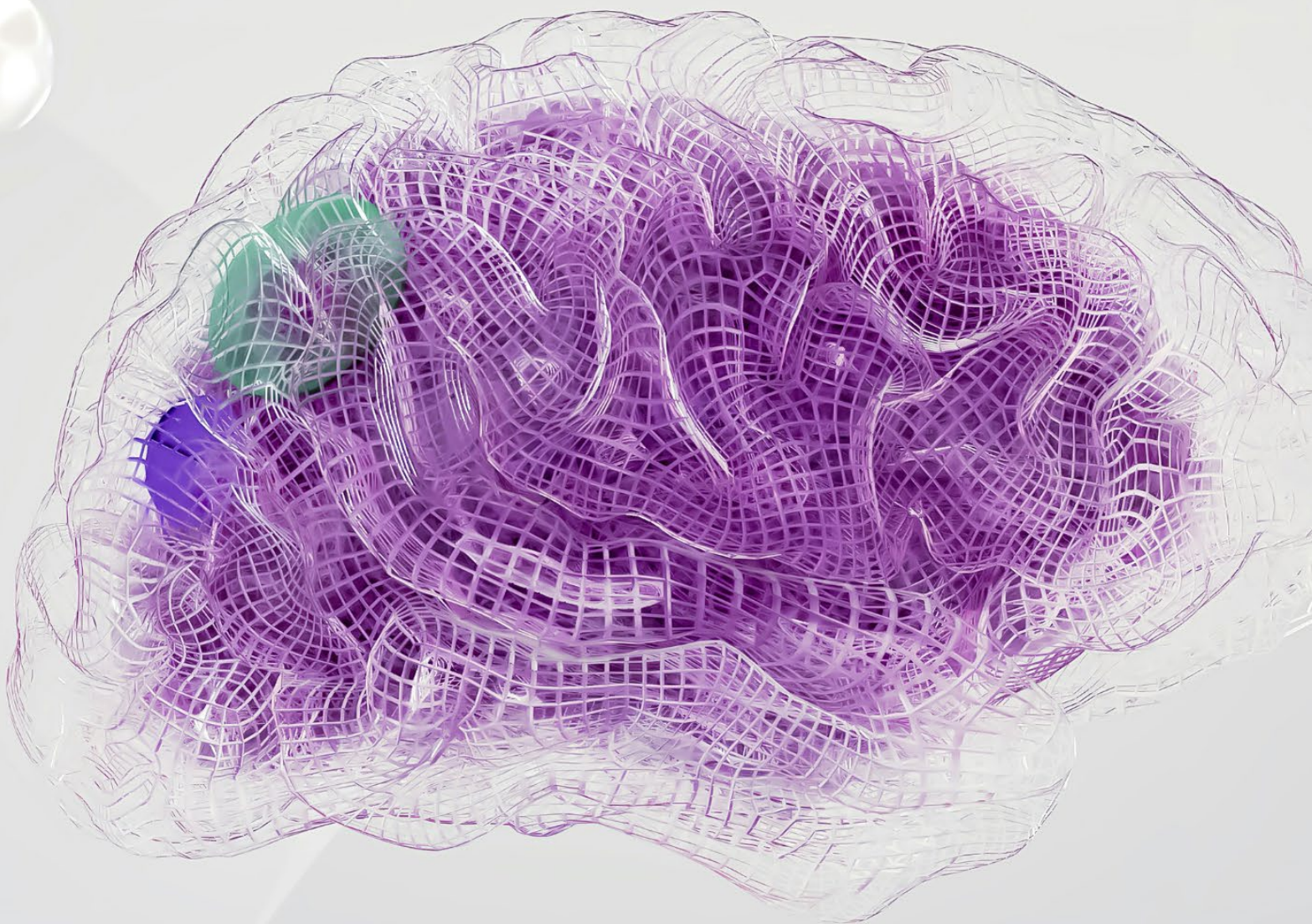


DES RECOMMANDATIONS SUR MESURE !

Comment les l'IA personnalisent votre expérience en ligne ?



INTRODUCTION

Dans un monde où nous sommes constamment submergés par une avalanche d'informations et de choix, comment pouvons-nous trouver ce qui nous intéresse vraiment ?

Comment Netflix sait-il quel film nous voudrions regarder parmi les 17 000 de sa sélection ?

Comment Spotify arrive-t-il à créer une playlist qui correspond parfaitement à nos goûts musicaux ?

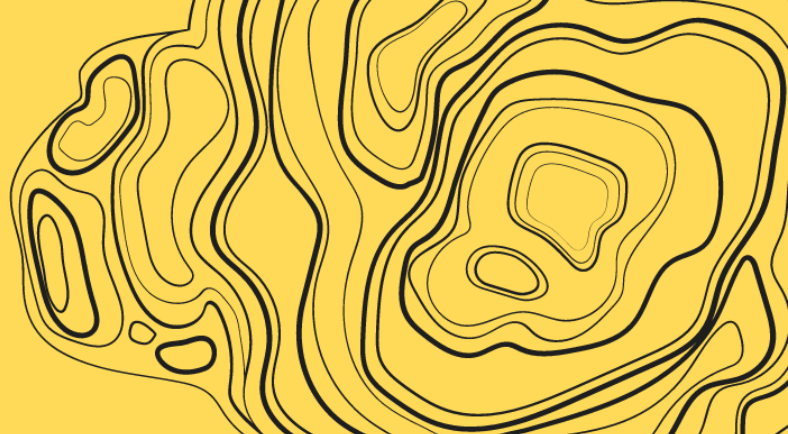
Comment Amazon devine-t-il quel livre pourrait nous plaire parmi les 410 000 disponible dans sa boutique Kindle ?

La réponse réside dans les algorithmes de recommandation, ces outils puissants qui façonnent notre expérience en ligne et nous aident à naviguer dans la montagne de contenu disponible.

Dans ce rapport sur les algorithmes de recommandation, nous explorerons leur fonctionnement, discuterons de leurs avantages et de leurs inconvénients, et découvrirons comment ils sont utilisés dans diverses applications en ligne. Nous examinerons également les défis éthiques et techniques qu'ils posent, et nous nous pencherons sur les tendances actuelles et futures dans ce domaine.

J'espère que ce voyage à travers le monde des algorithmes de recommandation sera aussi passionnant pour vous qu'il l'est pour moi. Alors, embarquons ensemble pour cette aventure à la découverte des algorithmes !

SOMMAIRE



INTRODUCTION

SOMMAIRE

GLOSSAIRE

Liste des notations

1

CHAPITRE 1

Compréhension et Objectifs	14
Types de données utilisées	16
Sur les traces des algorithmes de recommandation	18

2

CHAPITRE 2

Approche basé sur le contenu	20
Approche basé sur les utilisateurs	21
Approche hybride	24

3

CHAPITRE 3

Spotify	26
IKEA	30
TikTok	32



4

CHAPITRE 4

Reglementations	36
Mes données comme exemples	37

5

CONCLUSION

3

BILAN

Retour sur mes pistes de recherches	44
Projection sur les applications en entreprise	44
Ouverture sur les problématiques	44
Plan de diffusion	45
Réflexion personnelle	45

3

BIBLIOGRAPHIE

GLOSSAIRE

votre guide simplifié pour mieux comprendre

Système de recommandation (SdR)

Système de filtrage de l'information visant à présenter les éléments d'information qui sont susceptibles d'intéresser l'utilisateur.

Algorithmes de recommandation

Un système de recommandation est un algorithme d'intelligence artificielle ou d'IA, généralement associé au machine learning, qui utilise le Big Data pour suggérer ou recommander des produits supplémentaires aux consommateurs.

Filtrage collaboratif ou Collaborative Filtering (CF)

Technique de recommandation basée sur le comportement passé d'un utilisateurs similaires.

Filtrage basé sur le contenu ou Content-Based Filtering(CBF)

Technique de recommandation qui utilise les caractéristiques du contenu pour suggérer des éléments similaires à ceux que l'utilisateur a aimés dans le passé.

Filtrage collaboratif basé sur le voisinage ou Neighborhood-Based Collaborative

Filtering (NBCF)

méthode de recommandation de systèmes qui détermine l'intérêt d'un utilisateur pour un élément en se basant sur d'autres utilisateurs qui partagent certaines caractéristiques communes avec l'utilisateur.

Systèmes de recommandation hybrides ou Hybrid Recommender Systems

Systèmes qui combinent plusieurs techniques de recommandation, souvent en intégrant à la fois le filtrage collaboratif et le filtrage basé sur le contenu. L'objectif est d'améliorer la précision et la robustesse du système.

Algorithmes

Ensemble d'instructions ou une suite d'opérations qui permettra d'obtenir un résultat.

Big data

Somme colossale des données numériques provenant de nombreuses sources numériques, dont le Web, les réseaux sociaux, les bases de données privées, publiques à caractère commercial ou scientifique.

Intelligence artificielle

Demandé à un ordinateur de résoudre une tâche sans lui dire comment.

Machine Learning

Processus d'apprentissage des algorithmes, dont les conditions et les règles sont générées automatiquement par apprentissage statistique.

Deep Learning

Ensemble de méthodes d'apprentissage automatique tentant de modéliser, avec un haut niveau d'abstraction, des données.

Problème Démarrage à Froid

Situation dans laquelle le système de recommandation a du mal à fournir des suggestions précises pour les nouveaux utilisateurs ou les nouveaux éléments qui ont peu ou pas d'historique

Surpersonnalisation

Risque associé à trop de personnalisation, dans lequel le système de recommandation ne limite l'utilisateur à une bulle d'informations trop étroite, en ne lui montrant que des éléments similaires à ses préférences passées.

Ethique

Principes moraux qui guident le comportement humain. Dans le contexte de la technologie, l'éthique examine les implications morales des actions et des décisions liées à la technologie.

Tendances

référence aux mouvements populaires, aux styles ou aux comportements qui gagnent en popularité dans une société donnée à un moment précis.

Interactivité

Capacité d'un système à répondre aux actions des utilisateurs de manière dynamique, favorisant l'engagement et la participation.

Minicookies

Petits fichiers texte stockés sur votre ordinateur par votre navigateur Web lorsque vous visitez un site Web. Ils contiennent des informations telles que vos préférences de navigation, votre nom d'utilisateur et votre mot de passe, et sont utilisés pour améliorer votre expérience utilisateur sur le site Web.

Watch time

Le temps de visionnage est le temps total que les utilisateurs passent à regarder une vidéo sur une plateforme de partage de vidéos.

Micro-contenu

un type de contenu qui est court et facile à consommer. Il est souvent utilisé sur les réseaux sociaux pour attirer l'attention des utilisateurs et les inciter à interagir avec le contenu.

Métadonnées

Données qui décrivent d'autres données. Elles fournissent des informations sur les caractéristiques d'un ensemble de données, telles que la date de création, l'auteur, le format, la taille, etc.

Matrice

Tableau rectangulaire de nombres, de variables ou d'éléments. Les matrices sont utilisées dans de nombreux domaines, tels que les mathématiques, la physique, l'informatique, la statistique, etc.

Arbre de décision

Modèle de classification qui utilise une structure arborescente pour représenter les décisions et les conséquences possibles d'un ensemble de conditions.

Machine à vecteurs de support (SVM)

Algorithme d'apprentissage automatique qui est utilisé pour résoudre des problèmes de classification et de régression. Elle est basée sur la recherche d'un hyperplan qui sépare les données en deux classes.

Réseaux de neurones

Modèles d'apprentissage automatique qui sont inspirés du fonctionnement du cerveau humain. Ils sont utilisés pour résoudre des problèmes de classification, de régression, de reconnaissance de formes, de traitement du langage naturel, etc.

Règles d'association

Relations statistiques entre des variables dans un ensemble de données. Elles sont souvent utilisées pour découvrir des modèles dans les données, telles que les habitudes d'achat, les préférences des clients, les tendances de consommation, etc.

Web 1.0

Première phase du World Wide Web, caractérisée par des sites web principalement statiques, une diffusion d'informations unidirectionnelle et une interaction limitée.

Web 2.0

Evolution du Web vers des plateformes plus interactives, collaboratives et axées sur l'utilisateur, incluant les médias sociaux, les blogs, etc.

Netflix

Plateforme de streaming de contenu vidéo en ligne. Fondée en 1997, elle a commencé en tant que service de location de DVD par correspondance avant de devenir l'un des principaux fournisseurs de contenu en streaming, offrant des films, des séries télévisées, des documentaires et du contenu original.

Cinematch

Système de recommandation de films développé par Netflix pour suggérer des titres en fonction des préférences des utilisateurs.

Amazon

Entreprise américaine de commerce électronique fondée en 1994 par Jeff Bezos. Elle a débuté en tant que librairie en ligne avant de se diversifier pour devenir l'un des plus grands détaillants en ligne du monde, proposant une vaste gamme de produits, y compris des biens de consommation, des appareils électroniques et des services cloud.

Spotify

Plateforme de streaming musical fondée en 2006. Les utilisateurs peuvent accéder à un vaste catalogue musical, créer des listes de lecture, découvrir de nouveaux artistes et partager de la musique avec d'autres utilisateurs.

Flickr

Service de partage de photos en ligne lancé en 2004. Il permet aux utilisateurs de télécharger, organiser et partager leurs photos avec d'autres. Flickr a été particulièrement populaire pour ses fonctionnalités de partage de photos haute résolution.

Youtube

Plateforme de partage de vidéos en ligne fondée en 2005 et acquise par Google en 2006. Les utilisateurs peuvent télécharger, regarder, commenter et partager des vidéos sur une variété de sujets.

Friendster

L'un des premiers réseaux sociaux, lancé en 2002. Il permettait aux utilisateurs de se connecter avec des amis, de partager des photos et des messages, mais a perdu de sa popularité face à la concurrence.

MySpace

Réseau social populaire au début des années 2000, offrant des profils personnalisables, des fonctionnalités musicales et la possibilité de se connecter avec d'autres utilisateurs.

Usenet

Système de discussion en ligne introduit dans les années 1980. Il permet aux utilisateurs de participer à des forums de discussion sur une variété de sujets.

ByteDance

Une entreprise chinoise de technologie Internet qui a été fondée en 2012. Elle est surtout connue pour avoir créé l'application de partage de vidéos TikTok.

Toutiao

une application de médias sociaux chinoise qui fournit des informations personnalisées à ses utilisateurs. Elle utilise l'intelligence artificielle pour recommander des articles, des vidéos et d'autres contenus à ses utilisateurs.

Vine

une application de partage de vidéos qui permettait aux utilisateurs de créer et de partager des vidéos courtes. Elle a été lancée en 2013 et a été acquise par Twitter la même année. En 2017, Twitter a annoncé qu'elle allait fermer l'application.

Doyin

une application de partage de vidéos chinoise qui est similaire à TikTok. Elle a été créée par ByteDance et est disponible uniquement en Chine.

Tapestry .Inc

Entreprise multinationale américaine de mode de luxe. Elle est basée à New York et est la société mère de trois grandes marques : Coach New York, Kate Spade New York et Stuart Weitzman.

LISTE DES NOTATIONS

U ensemble des utilisateurs u .

I ensemble des items i .

I_u ensemble des items notés par l'utilisateur u .

$I_{u,v}$ ensemble des items notés par l'utilisateur u et v .

r rating.

$r_{u,i}$ représente la note que l'utilisateur u a assigné explicitement à l'item i .

\hat{r} rating prédit.

$\hat{r}_{u,i}$ représente la prédiction de la note que l'utilisateur u assignerait à l'item i .

CHAPITRE 1

Comprendre les algorithmes de recommandation

Les algorithmes de recommandation sont omniprésents dans notre vie quotidienne numérique, guidant discrètement nos choix en ligne. Que ce soit pour choisir un film à regarder, un livre à lire, ou même une nouvelle recette à essayer, ces algorithmes jouent un rôle clé.

Dans ce chapitre, nous allons plonger dans l'univers de ces algorithmes. Nous allons remonter le temps pour découvrir comment ils ont évolué au fil des ans, comprendre les mécanismes et explorer les différentes approches utilisées pour prédire ce que nous pourrions aimer ou trouver intéressant.

Compréhension et Objectifs des Algorithmes de Recommandation

Imaginez que vous êtes dans une immense bibliothèque, remplie de millions de livres. Comment trouveriez-vous le livre parfait pour vous ? Les systèmes de recommandation sont comme des guides qui vous aident à trouver le livre idéal en fonction de vos préférences et de vos intérêts passés.

Un système de recommandation est en fait un algorithme d'intelligence artificielle, généralement associé à l'apprentissage automatique, également appelé Deep Learning, qui utilise de grandes quantités de données pour suggérer ou recommander des produits supplémentaires aux consommateurs. Cette grande quantité de données pose également quelques problèmes notamment éthique. Ces données peuvent être de différentes natures, tels que les achats passés, l'historique de recherche, les informations démographiques mais également des données d'interaction comme des impressions, des commentaires, ...

Les systèmes de recommandation sont donc entraînés à comprendre les préférences, les décisions et les caractéristiques des personnes et des produits en utilisant des données recueillies sur leurs interactions précédentes.

Mais quels sont les impacts de ses systèmes de recommandation sur votre activité numérique et sur les entreprises qui les utilisent.

Pour ce qui est des avantages pour l'utilisateur final on retrouve principalement le développement de l'expérience utilisateur grâce à la personnalisation.

Mais ces algorithmes ont également quelques impact négatifs. Comme évoqué plutôt, les risques liés à la protection de la vie privée et l'éthique. De nos jours les utilisateurs finaux de ses services numérique sont de plus en plus avertie et strict sur la récupération de leurs données. Deux autres risques de ses systèmes qui peut nuire à l'expérience utilisateur sont :

- Le **démarrage à froid**, ce qui correspond à la situation dans laquelle le système de recommandation a du mal à fournir des suggestions précises pour les nouveaux utilisateurs ou les nouveaux éléments qui ont peu ou pas d'historique, ce qui peut engendrer des recommandations erroné.
- La **surpersonnalisation**, un phénomène qui vient enfermer l'utilisateur dans une bulle d'informations trop étroite.

Les algorithmes de recommandation ont également des impacts sur les fournisseurs de ces services. Parmi les impacts positifs, on peut lister :

- L'augmentation du taux de conversion et des revenus
- La fidélisation des clients
- L'optimisation de la sélection et de la diffusion du contenu

Pour vous donner des idées de l'impact de ces technologies, voici quelques chiffres :



Figure 1- Chiffres sur l'impact des algorithmes de recommandation

Source : [IntoTheMinds](#)

Types de Données Utilisées dans les Algorithmes de Recommandation

La principale ressource nécessaire au fonctionnement de ces systèmes de recommandation, c'est la donnée. Sans données, aucun algorithme ne sera en capacité de vous recommander du contenu si il ne dispose d'aucune données. L'algorithme se base sur 3 éléments de bases :

- Un **ensemble d'utilisateurs** : $u \in U$
- Un **catalogue d'items** : $i \in I$
- Une collection de **retours d'utilisateurs** (feedback)

Ces données si nécessaire sont également appelé feedback, il en existe 2 types :

- Les feedbacks explicites

Présenter sous la forme de notes (ratings en anglais) que les utilisateurs attribuent aux items. Les ratings sont souvent saisis sur une échelle graduée qui indique le niveau d'appréciation de l'utilisateur par rapport aux items concernés. Il permettent de quantifier ce que l'on aime ou ce que l'on n'aime pas. Un exemple typique, largement utilisé dans de nombreux systèmes de recommandation populaires tel que Amazon, Netflix, ... consiste à attribuer les ratings sur l'ensemble $\{1, 2, 3, 4, 5\}$. Dans ce contexte, un rating de 1 indique une insatisfaction profonde alors que 5 indique l'appréciation maximale.



Figure 2 - Exemple de feedback explicite (Source Amazon)



Le système de notation par **étoiles**, omniprésent aujourd'hui, a vu le jour chez Amazon en 1995 ou 1996. Il avait pour but d'inciter les utilisateurs à noter les produits. Amazon introduit donc une fonctionnalité d'évaluation en un clic, sans avoir à remplir un formulaire complet.

- Les feedback implicites

Collecté à la place, ou en complément des retours explicites, un système de recommandation peut aussi disposer de retours implicites. Il s'agit souvent de traces basées sur les comportements des utilisateurs ou les interactions « utilisateur-item », qui permettent de montrer implicitement leur intérêt par rapport aux items avec lesquels ils ont interagit. Par exemple, si un utilisateur écoute fréquemment la musique d'un artiste, il est raisonnable de supposer qu'il aime cet artiste. De même, si un utilisateur a acheté un vêtement et ne le retourne pas, on peut supposer que c'est parce qu'il l'apprécie. Bien qu'on ne dispose pas de retour explicite de l'utilisateur pour l'artiste ou le vêtement en question, des hypothèses peuvent être posées. On le représente donc plus de manière binaire $\{0$ ou $1\}$. Ces retours implicites peuvent être regroupés dans différentes catégories :

- **Attention**

- Le défilement de la page

En analysant la vitesse de défilement, on peut voir si l'utilisateur a porté de l'importance à des éléments présent sur la page.

- Le suivi du curseur

Le suivi du curseur peut indiquer si l'utilisateur a vu les recommandations. Des études montrent une corrélation entre les mouvements des yeux et du curseur, permettant de savoir si une recommandation a été vue par l'utilisateur.

- L'aspects tactiles (sur mobile essentiellement)

Un zoom sur une recommandation est un moyen évident de conclure sur l'attention de l'utilisateur.

- **Interactions**

- Le temps de consommation

Le temps de consommation est interprété différemment selon la plate-forme. Par exemple, sur Facebook, regarder une vidéo pendant plus de 3 secondes est considéré comme une vue, tandis que sur YouTube, il faut regarder une vidéo pendant au moins 30 secondes pour qu'elle soit considérée comme une vue. C'est ce qu'on appelle le watch-time.

- L'approfondissement

Lorsque l'utilisateur pousse la recherche, ça témoigne d'un intérêt pour le contenu. On retrouve cette même mécanique, dans l'interface de Spotify vous pouvez cliquer sur le nom de l'artiste pour obtenir plus d'informations à son sujet.

- L'enregistrement d'un contenu (une interaction forte parmi d'autres)

Enregistrer un élément dans une liste de lecture est un signal fort qui indique que vous appréciez la recommandation.

- La re-consommation

La re-consommation d'un article est un signal extrêmement fort, même si ce comportement n'est pas nécessairement lié à un sentiment positif.

Une des manières de représenter ces feedback est d'utiliser une matrice dite de rating. On obtient donc une matrice avec les lignes qui représente les utilisateurs et les colonnes représentent les items. La figure suivante illustre un exemple de deux matrices de ratings qui représentent les retours de 4 utilisateurs sur 4 items.

	i_1	i_2	i_3	i_4
u_1	1			5
u_2		3		
u_3	2		4	1
u_4			2	

	i_1	i_2	i_3	i_4
u_1	1			1
u_2		1		
u_3	1		1	1
u_4			1	

Figure 3 - Exemple de matrice de rating (explicite à gauche, implicite à droite)

Sur les Traces des Algorithmes de Recommandation

Pour la première apparition des algorithmes de recommandation, il faut revenir à l'émergence du web 1.0. Pour rappel, le Web 1.0 a vu le jour dans les années 1990 et représente la première phase du web. Inventé par Sir Tim Berners-Lee, un informaticien britannique, le web était principalement un réseau de diffusion de contenu. Les pages étaient statiques, sans interactivité, et le contenu était à sens unique. Le but principal du Web 1.0 était de connecter les informations et de les rendre accessibles à tous.

C'est donc en 1992 que Paul Resnick et son collègue John Riedl, tous deux chercheurs en informatique ont proposé le premier système de recommandation pour les articles du forum Usenet. Il s'agit d'un ancien système de discussion en ligne qui a émergé dans les premières années d'Internet, bien avant l'avènement des médias sociaux tels que nous les connaissons aujourd'hui. On pouvait y accéder depuis des lecteurs de nouvelles, il a récemment été mis de côté par Google qui ne le supporte plus sur ses services. Usenet n'est pas techniquement considéré comme un réseau social mais il était parmi les précurseurs à cette époque. En prenant en compte le fait que les contenus étaient à principal but informatif, c'est donc un algorithme de recommandation basé sur le contenu (Content-Based Filtering) ou « item-to-item » mais on y reviendra plus tard.

Il faudra ensuite attendre l'arrivée du Web 2.0 dans les années 2000, pour voir l'apparition du filtrage collaboratif (Collaborative Filtering), ou « user-to-user ». On y voit un développement de plusieurs choses, notamment :

- Interactivité, les utilisateurs pouvaient désormais participer activement aux contenus.
- Changement de stratégie et l'émergence de plateformes de partage, Netflix (1997 et virage numérique en 2007), Flickr (2004), YouTube (2005)
- Arrivée des premiers réseaux sociaux, Friendster (2002), MySpace (2003), Facebook (2004), ...
- Introduction de la publicité ciblée avec le système des minicookies.



En 2006, **Netflix** a organisé un concours nommé Netflix Prize pour améliorer son algorithme de recommandation de films et de séries. Le concours visait à améliorer leurs algorithmes de l'époque Cinematch, dans le but de recommander des contenus aux utilisateurs en fonction de leurs préférences.

Le grand prix, d'une valeur de 1 million de dollars, a été remporté par l'équipe de BellKor's Pragmatic Chaos, qui a réussi à améliorer l'algorithme de Netflix de 10,06%.

Le dernier type d'algorithme de recommandation appelé filtrage hybride n'a pas de date précise de création ou d'arrivée sur une application. Cependant, il se développe largement car, comme son nom l'indique, il est hybride, ce qui signifie qu'il prend les meilleurs points des deux systèmes arrivés plus tôt. Aujourd'hui, c'est l'un des plus répandus.

Chacun de ses 3 types de systèmes de recommandation a ses avantages et ses inconvénients et ces axes seront développés dans le chapitre 2.

CHAPITRE 2

Quel algorithme pour quel application

Comme présenté dans la partie historique des algorithmes de recommandation, il existe différents types ayant chacun ses particularités notamment au niveau des données utilisés pour les recommandations.

Une approche basé sur les items, le filtrage basé sur le contenu (CBF)

Le filtrage basé sur le contenu est une technique de filtrage qui se base sur les caractéristiques intrinsèques des items pour recommander des articles similaires à l'utilisateur. Contrairement au filtrage collaboratif, le filtrage basé sur le contenu ne se base pas sur les opinions des autres utilisateurs. Au lieu de cela, il utilise des informations sur les articles eux-mêmes pour recommander des articles similaires.

Par exemple, si un utilisateur a acheté un livre sur la programmation informatique, le système de recommandation peut recommander d'autres livres sur la programmation informatique en se basant sur les caractéristiques de ce livre, telles que le sujet, l'auteur, le style d'écriture, etc. Le filtrage basé sur le contenu est souvent utilisé dans les systèmes de recommandation de musique et de films, où les caractéristiques des articles sont facilement accessibles.

Cette approche se décompose en différentes étapes :

- **Pré-traitement en analyse**, extraction des caractéristiques pertinentes des items, telles que le genre, l'auteur, le sujet, etc., et à les représenter sous forme de vecteurs.
- **Construction des profils des utilisateurs**, construction d'un profil pour chaque utilisateur en se basant sur les articles qu'il a consultés ou achetés dans le passé.
- **Filtrage, appariement et recommandation**, comparaison des profils des utilisateurs avec les caractéristiques des articles pour recommander des articles similaires.
 - **Modèle des k plus proches voisins**, rechercher les k articles les plus similaires à l'article consulté par l'utilisateur et à les recommander.
 - **Modèle de régression linéaire**, prédiction de la note que l'utilisateur donnerait à un article en se basant sur les notes qu'il a données à des articles similaires.
- **Comparaison avec les approches de filtrage collaboratif**, comparaison des résultats obtenus avec les approches de filtrage basé sur le contenu avec ceux obtenus avec les approches de filtrage collaboratif.

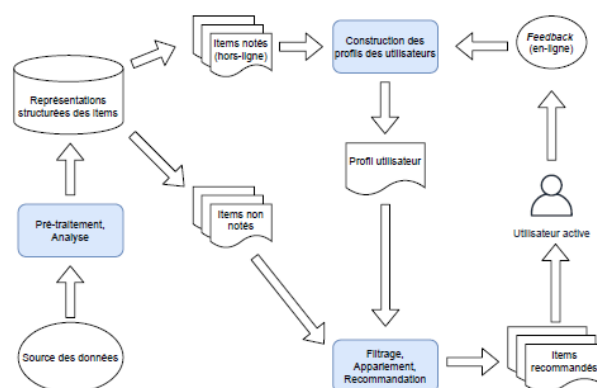


Figure 4 - Architecture générale des approches de recommandation basées sur le contenu (Source Thèse Yu Du)

Une approche basé sur les utilisateurs, le filtrage collaboratif (CF)

Le filtrage collaboratif, ou Collaborative Filtering (CF), est une technique couramment utilisée dans les systèmes de recommandation scientifiques et industriels.

Pour recommander des articles à l'utilisateur u , il s'appuie sur les opinions des autres utilisateurs, d'où le terme "collaboratif". GroupLens est l'un des premiers systèmes de recommandation à utiliser le filtrage collaboratif. Dans le domaine du commerce électronique, le système de recommandation d'Amazon utilise également cette technique, comme en témoigne le titre de la section "Les clients ayant acheté ce produit ont également acheté...". Dans un système de recommandation, il est rare que les utilisateurs donnent un avis (rating) pour chaque article du catalogue disponible, ce qui a pour conséquence d'obtenir une matrice des ratings incomplète.

L'objectif du filtrage collaboratif est donc de prédire les valeurs des ratings manquants en se basant sur les ratings disponibles.

L'approche du filtrage collaboratif est divisée en deux sous-catégories :

- l'approche filtrage collaboratif basée sur la mémoire
- l'approche filtrage collaboratif basée sur les modèles

L'approche basés sur la mémoire / le voisinage

L'approche de filtrage collaboratif basée sur la mémoire, également appelée filtrage collaboratif basé sur le voisinage ou Neighborhood-Based Collaborative Filtering (NBCF), est l'un des premiers algorithmes de filtrage collaboratif. Il se base sur 2 hypothèses :

1. Les utilisateurs similaires notent les articles de manière analogue.
2. Les articles similaires reçoivent des notes similaires.

Pour l'histoire, ce sont les développeurs de Tapestry qui ont été les premiers à utiliser le terme "filtrage collaboratif" comme moyen de recueillir des données qualitatives. Développé à Xerox PARC comme moyen de traiter les grandes quantités de courriels et de messages postés sur les groupes de discussion. La conception et le concept sous-jacents reposaient sur l'idée que le processus de filtrage serait plus efficace si l'on faisait appel à des êtres humains.

Dans cette sous-catégorie de l'approche collaborative on distingue encore deux approches :

- NBCF basé sur les utilisateurs : Ils recommandent des éléments à un utilisateur en se basant sur les préférences d'autres utilisateurs qui ont des goûts similaires.
- NBCF basé sur les items : ils recommandent des éléments à un utilisateur en se basant sur les caractéristiques des éléments eux-mêmes.

Les approches basés sur le voisinage (NBCF) sont constituées de 2 étapes clés consécutives, l'identification du voisinage et la prédiction des rating.

Pour faire court, cette première partie sur l'identification du voisinage. Elle consiste à identifier un ensemble de voisins pour chaque utilisateur en se basant sur leurs préférences et leurs comportements passés. Il utilise notamment des outils mathématique comme la corrélation de Pearson, la similarité cosinus, ...

Et pour ce qui est de la phase prédiction des rating, l'algorithme calcul une note pour un item en utilisant les notes des voisins de l'utilisateur pour cet item et en utilisant une technique de pondération pour donner plus de poids aux voisins les plus similaires.

Pour que vous compreniez bien le fonctionnement de cette méthode très répandu, voici un exemple concret (tiré de la thèse de Yu Du présenté en décembre 2021). On considère la matrice de ratings suivantes. L'objectif est de prédire la valeur de rating de l'utilisateur u_1 pour l'item i_2 . On nomme cette prédiction donc \hat{r}_{u_1, i_2} (case verte).

	i_1	i_2	i_3	i_4	i_5	i_6	i_7
u_1	1		3	1			2
u_2	2	1		5	5		
u_3		2	3				
u_4	1	2					
u_5	2	5	3	1			1
u_6	1		3	1		4	2

Figure 5 - Exemple d'une matrice des ratings pour illustrer l'approche NBCE (Source Thèse Yu Du)

Avant de calculer le rating en lui-même, dans ce cas si on va sélectionner les utilisateurs a partir des quels calculer la similarité. Ici on va calculer la similarité de l'utilisateur u_1 et les utilisateurs u_2, u_3, u_4 et u_5 . Pour ce faire on utilise la formule du Coefficient de corrélation de Pearson :

$$sim_{PCC}(u, v) = \frac{\sum_{i \in I_{u,v}} [(r_{u,t} \times \bar{r}_u)(r_{v,t} - \bar{r}_v)]}{\sqrt{\sum_{i \in I_{u,v}} (r_{u,t} - \bar{r}_u)^2 \sum_{i \in I_{u,v}} (r_{v,t} - \bar{r}_v)^2}}$$

$$sim_{PCC}(u_1, u_2) = \frac{\sum_{i \in I_{u,v}} [(r_{u,t} \times \bar{r}_u)(r_{v,t} - \bar{r}_v)]}{\sqrt{\sum_{i \in I_{u,v}} (r_{u,t} - \bar{r}_u)^2 \sum_{i \in I_{u,v}} (r_{v,t} - \bar{r}_v)^2}}$$

$$= \frac{(1 - 1,75)(2 - 3,25) + (1 - 1,75)(5 - 3,25)}{\sqrt{[(1 - 1,75)^2 + (1 - 1,75)^2][(2 - 3,25)^2 + (5 - 3,25)^2]}}$$

$$= -0,16$$

On applique la même méthode pour les autres valeurs de similarité : $sim_{PCC}(u_1, u_3) = 1,00$; $sim_{PCC}(u_1, u_4) = 1,00$; $sim_{PCC}(u_1, u_5) = 0,50$

Une fois ces similarité calculé, on passe à la deuxième étape : la prédiction. Il faut tout de même noté que cette exemple est à petite échelle, faible nombre d'utilisateur et d'items. On applique maintenant la formule de prédiction :

$$\hat{r}_{u,i} = \bar{r}_u + \frac{\sum_{v \in N_{u(i)}} [(r_{v,i} - \bar{r}_v) \times \text{sim}(u, v)]}{\sum_{v \in N_{u(i)}} \text{sim}(u, v)}$$

$$\hat{r}_{u_1, i_2} = \bar{r}_{u_1} + \frac{\sum_{v \in N_{u_1(t_2)}} [(r_{v,t} - \bar{r}_v) \times \text{sim}(u_1, v)]}{\sum_{v \in N_{u_1(t_2)}} \text{sim}(u_1, v)}$$

$$= 1,75 + \frac{(2 - 2,5) \times 1,00 + (2 - 1,5) \times 1,00 + (5 - 2,4) \times 0,50}{1,00 + 1,00 + 0,50}$$

$$= 2,27$$

Ici, on peut noter que le résultat n'est pas entier comme les autres valeurs de la matrices, mais elle correspond à un rating prédit et non réel.

L'approche basés sur les modèles

Les approches NBCF sont des méthodes spécifiques aux instances, où une instance est la valeur de $\hat{r}_{u,i}$ à prédire. En revanche, il existe des approches de CF basées sur les modèles dans lesquelles il y a une distinction claire entre la phase d'entraînement et la phase de prédiction.

Dans le cas des approches NBCF, un modèle n'est pas spécifiquement créé à l'avance pour la prédiction, autre qu'une phase de pré-traitement (par exemple, calculs des similarités, calculs des prédictions basées sur des équations pré-définies, ...).

En revanche, pour les approches basées sur les modèles, un modèle d'apprentissage automatique (supervisé en général) est d'abord créé. Les paramètres de ce modèle sont ajustés sur la base des ratings disponibles dans le jeu de données, en optimisant une fonction objectif. Cette étape est appelée entraînement du modèle. J'espère que cela répond à votre question. N'hésitez pas à me demander si vous avez besoin de plus d'informations.

Les modèles d'apprentissage automatique tels que les arbres de décision, les machines à vecteurs de support (SVM) ou les réseaux de neurones ont été largement utilisés pour résoudre des problèmes pratiques tels que la classification et la régression. Ces modèles peuvent tous être adaptés pour résoudre le problème de la prédiction des ratings, c'est-à-dire la complétion de la matrice de ratings.

En effet, on peut considérer la prédiction des ratings comme une généralisation des autres problématiques évoquées. Dans un cadre de classification, par exemple, il s'agit également de compléter une matrice, dans laquelle tous les champs sauf ceux de la dernière colonne (la variable de classe) sont remplis par des valeurs spécifiques (c'est-à-dire les caractéristiques des jeux de données).

Plus généralement, dans le cas du filtrage collaboratif, chaque case de la matrice dont on dispose peut être vide. Par conséquent, le problème de la prédiction des ratings peut être considéré comme une généralisation de la classification ou de la régression. De nombreux modèles ont été appliqués dans le contexte du filtrage collaboratif pour prédire les ratings. On peut lister :

- **Modèle baseline**, basé sur le biais, c'est un des modèles de référence.
- **Modèle SVD**, basé sur les facteurs latents. Il dispose également d'une extension SVD++ qui vient prendre en compte les interactions utilisateur-item (feedback implicite).
- **Modèle NeuMF** (Neural Matrix Factorization), mise en place du deeplearning.
- ...

Une approche mêlant les 2 précédents, le filtrage hybrides

Les approches hybrides dans les algorithmes de recommandation combinent plusieurs techniques de recommandation pour améliorer la qualité et la diversité des recommandations proposées. Les approches hybrides peuvent combiner des techniques de filtrage collaboratif et de filtrage basé sur le contenu.

Les approches hybrides sont souvent utilisées dans les systèmes de recommandation de commerce en ligne et de musique, où les utilisateurs ont des préférences et des comportements d'achat complexes. Les approches hybrides peuvent améliorer la qualité des recommandations en combinant les avantages de différentes techniques de recommandation. Les approches hybrides sont une solution pour pallier les inconvénients majeurs des approches de filtrage collaboratif et de filtrage basé sur le contenu.

Comme l'ensemble des approches présenté plus tôt, il existe différentes hybridation :

- **Les hybridations pondérées**, cette approche consiste à combiner les scores obtenus par différents systèmes de recommandation de manière linéaire. La prédiction finale du rating que l'utilisateur u donnerait à l'item i est une combinaison des prédictions faites par différents systèmes de recommandation. Les poids des différents systèmes de recommandation peuvent être ajustés en fonction des feedbacks des utilisateurs en ligne.
- **Les hybridations commutées**, le système bascule entre différentes techniques de recommandation selon certains critères. Par exemple, une approche CBF est utilisée lorsque le nombre des ratings est petit, tandis qu'une approche CF est utilisée lorsque le système dispose de suffisamment de retours.
- **Les hybridations mixtes**, présenter une liste d'items qui contient des recommandations de chaque approche utilisée. Les items résultants de chaque approche de recommandation sont présenté simultanément à l'utilisateur.
- **Les hybridations en cascade**, cette approche consiste à utiliser plusieurs systèmes de recommandation en cascade. Une première technique de recommandation est utilisée pour générer un classement grossier des items candidats, puis une seconde technique permet d'affiner cette première liste de candidats, et ainsi de suite.
Notons que cette hybridation est sensible à l'ordre des étapes : un CF-CBF produirait certainement des résultats différents d'un CBF-CF.

Les algorithmes de recommandation sont un domaine en constante évolution, avec de nombreux acteurs investis dans la recherche pour améliorer leurs efficacité et leurs précision. Ils sont basés sur le contenu, sur les utilisateurs et une version hybrides ont tous leurs avantages et leurs inconvénients, et leurs choix dépend des données disponibles et des objectifs du service. Ces algorithmes complexes sont utilisés dans de nombreux domaines, tels que la musique, les films, les produits, les publicités et les réseaux sociaux, pour n'en citer que quelques-uns mais cette liste va sûrement continuer de grossir dans les années qui arrivent.

CHAPITRE 3

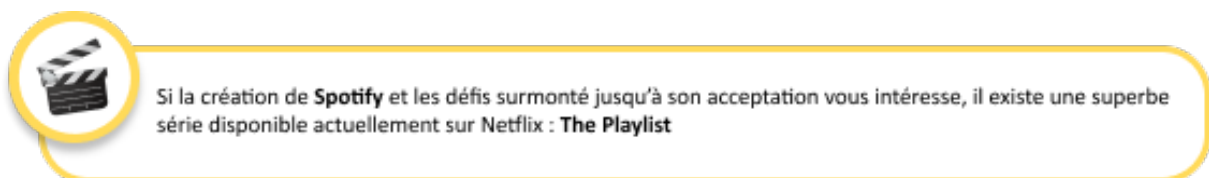
Algorithmes de recommandation : Diverses applications pour améliorer l'expérience utilisateur

Comme vous l'avez sûrement compris, les données représentent le cœur des systèmes de recommandation, avec la recherche de pousser toujours plus proche des attentes des utilisateurs, les plateformes sont de plus en plus intrusives. Dans ce dernier chapitre, nous allons aborder les différentes solutions qui visent à réguler la collecte de ces données, quelques cas concrets d'application de ces algorithmes (Spotify, IKEA, ...) et pour finir un exemple concret des données que des services grand public comme Instagram ou TikTok récupèrent pour alimenter leurs algorithmes de recommandation.

Les algorithmes sont devenus un élément clé de la stratégie commerciale de nombreuses entreprises. Peu importe le domaine, toutes se tournent vers la personnalisation pour maximiser leurs impacts, y compris leurs ventes. Des géants tels que Ikea, Spotify ou encore TikTok ne dérogent pas à la règle.

Le cas de Spotify

Pour sa part Spotify, n'a pas vraiment pu choisir entre filtrage basé sur le contenu et collaboratif, il fait cohabiter les deux approches dans son service de musique lancé en 2008 et fondé par Daniel Ek et Martin Lorentzon.



Durant les années qui suivent son lancement, Spotify a dû lutter pour être accepté, notamment sur :

- Les droits musicaux
- Numérisation de la musique (baisse des ventes physiques)
- La sécurité des données des utilisateurs compromise par des cyberattaques

Il faudra attendre 2014 et le rachat de la société Echo Nest, une plateforme de données et d'intelligence musicale pour voir arriver les premiers systèmes de recommandations sur la plateforme. L'algorithme de Spotify est directement visible à plusieurs endroits sur l'application.

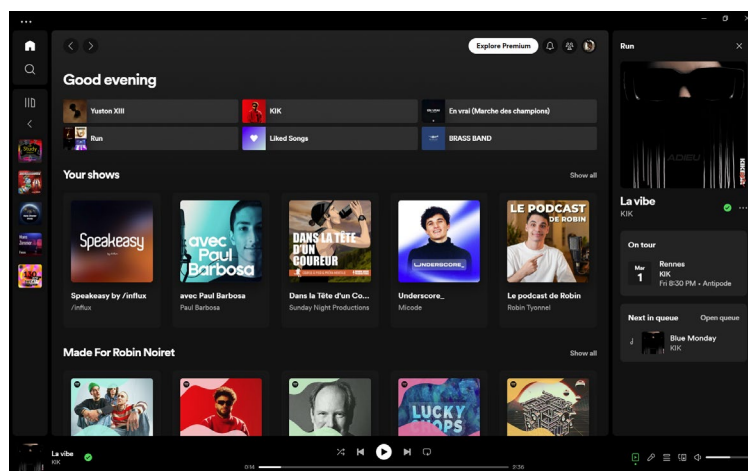


Figure 6 - Exemple de page d'accueil Spotify

Déjà sur la page d'accueil on retrouve une section 'Made for' qui vient proposer différentes playlist composé de musique que l'utilisateur a déjà aimé ou qu'il pourrait aimer (musique d'un artiste suivie, ...) C'est une première mise en avant des recommandations. On peut voir également 2 autres exemples de recommandation qui illustrent l'utilisation des systèmes :

- Basé sur le contenu
- Basé sur les utilisateurs

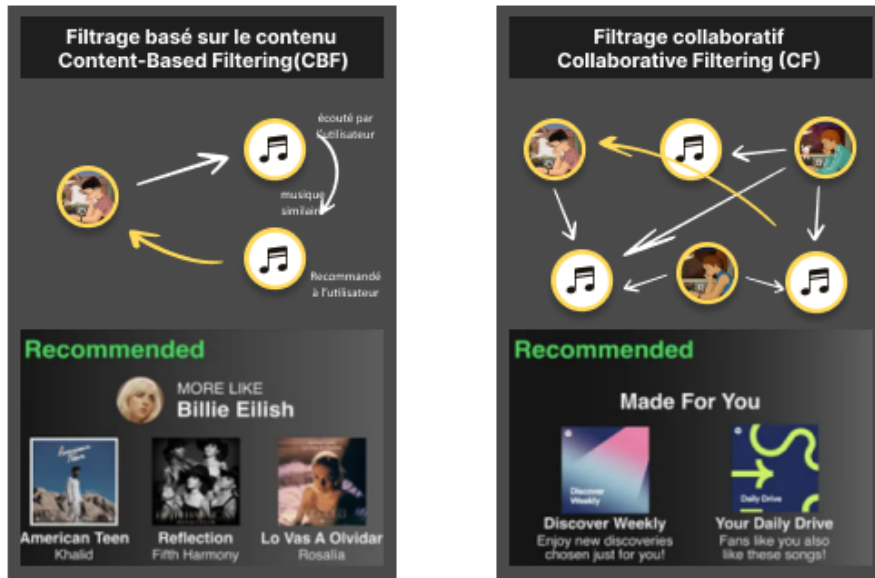


Figure 7 - Exemple utilisation des différentes approches dans Spotify

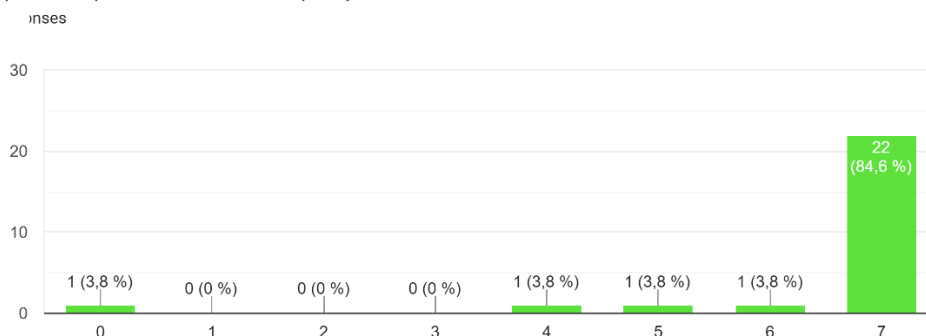
Mais cette utilisation du système basé sur le contenu (CBF) chez Spotify peut également engendrer une surpersonnalisation de vos recommandations et vous enfermer dans une bulle.

Sondage pour les utilisateurs de Spotify

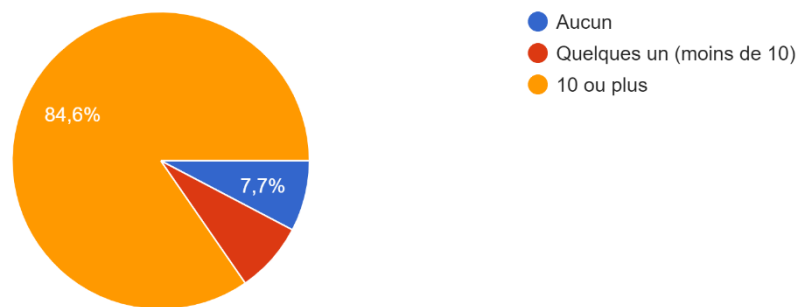
Dans le cadre de cette veille technologique, j'ai créé un sondage pour regrouper les avis venant de divers horizons. Cette étude n'a rien de très pointu, ce sont des valeurs à but d'illustrer des problèmes identifiés des algorithmes. Cette étude regroupe 27 interrogés mixtes dans une tranche d'âge en [17 et 30 ans].

Parmi les interrogés, 92,6% utilise Spotify et c'est sur ces personnes que se basent le reste de l'étude.

À quelle fréquence utilisez-vous Spotify ?



Combien de chansons avez vous découvert grâce à l'algorithme de recommandation de Spotify ?



Les prochaines questions moins orientées chiffres, mais plutôt vers expérience personnelle des interrogés, je vais citer quelques réponses.

Dans quelle mesure l'algorithme de Spotify influence-t-il vos choix musicaux au quotidien ?

« Si vous écoutez que des chansons tristes on se retrouve dans une boucle de chanson triste. »

« il reste dans les style que j'écoute. »

« Il m'influence beaucoup grâce à la fonctionnalité aléatoire amélioré. »

« assez souvent (radar/découvertes). »

« J'utilise régulièrement les playlists personnalisées de Spotify pour découvrir de nouvelles musiques susceptibles de me plaire. »

Quelles sont vos attentes principales lorsque vous utilisez l'algorithme de recommandation de Spotify ? (Découvrir de nouveaux artistes/musiques, ...)

« Découvrir de nouveaux artistes, de nouveaux styles »

« Découvrir de nouvelle musique et de nouveaux podcast »

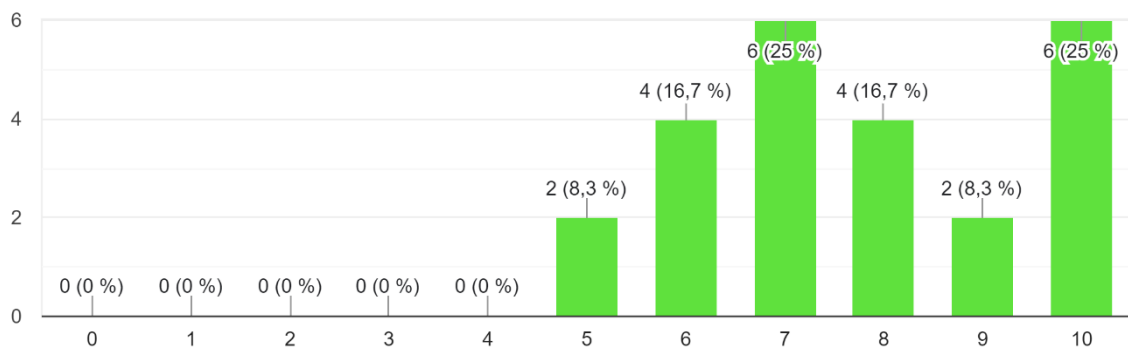
« Découvrir des musiques en rapport avec mes goût actuel. »

« Avoir une variété de musique pour éviter les redondances de ma playlist ou avoir entièrement des musiques suggérer qui colleront à mon style d'écoute »

« Qu'ils identifient au mieux mes goûts, tout en me faisant découvrir de nouveaux genres, styles ou artistes musicaux. »

« Je m'attends à avoir des recommandations très intéressantes (des artistes dans les styles que j'écoute et des musiques qui ressemblent à celles ajoutés dans ma playlist). »

Selon votre expérience, estimez-vous que les recommandations actuelles correspondent à vos goûts personnels ?



Pour finir je trouvais ça intéressant d'avoir l'avis d'utilisateur sur les améliorations qu'ils apporteraient à cet algorithme.

« Plus nous sortir un peu de notre zone de confort. »

« qu'on puissent plus facilement enlever cette fonctionnalité. »

« Pour moi il n'y a rien à améliorer, il est déjà très bien pour mon utilisation. »

« Pour apporter une plus grande pertinence, on pourrait penser mettre de nouvelles recommandations par rapport au nombre d'artistes ajoutés, au nombre d'albums ajoutés/téléchargés de ces artistes et aux titres ajoutés de ces artistes (en veillant aussi au style de chaque titre), suite aux recommandations actuelles. »

« Plus de recommandations et moins d'écoute des musiques qui sont déjà dans mes playlist. »

Comme le témoignent certaines des réponses tirées du sondage présenté ci-dessus, la surpersonnalisation dans ce genre de services peut avoir des effets négatifs. En effet, certains algorithmes de recommandation comme celui de Spotify ont tendance à limiter la diversité des genres et des artistes proposés. Cela peut également conduire à une standardisation de la musique, où seuls les artistes les plus populaires sont écoutés.

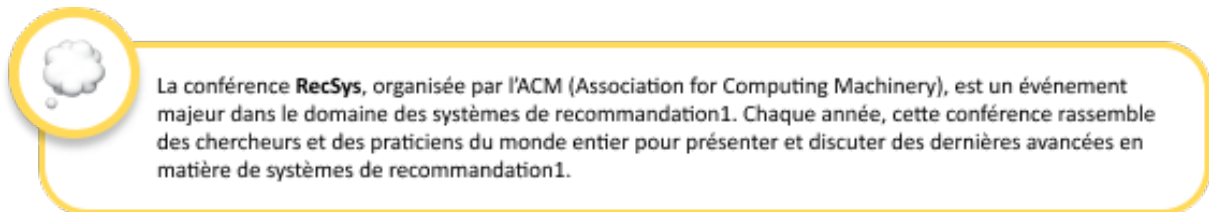
Dans le cas d'un système basé sur le contenu, l'algorithme va donc pousser des musiques et des artistes dans le même style que les habitudes de consommations de l'utilisateur. Et dans le cas d'un système basé sur les utilisateurs, prenant en compte la grande quantité d'utilisateur, 574 millions d'utilisateurs actifs mensuels à la fin du troisième trimestre 2023 d'après les chiffres officiels. En prenant en compte la domination de certains styles comme le Hip Hop, (38%), Pop (27%), ... L'algorithme va naturellement rediriger les utilisateurs vers les styles répandus.

Pour reprendre leurs mots, les défenseurs de l'industrie musicale traditionnelle disent : « utiliser des algorithmes personnalisés pour la recommandation musicale, c'est déshumaniser la découverte musicale. » Selon eux, remettre nos oreilles entre les mains d'un programmeur radio ou d'un disquaire pourrait nous mener vers des rythmes insoupçonnés, à l'inverse des algorithmes qui se plient à ce que nous aimons déjà, détruisant toute notion d'aléatoire.

Les algorithmes sont donc bénéfique à la plateforme en venant améliorer l'expérience utilisateur, mais il peut également avoir des impacts limitant. Cependant les algorithmes comme celui de Spotify ou autre sont en constante évolution pour se perfectionner.

Le cas de Ikea

L'exemple de Spotify tout comme celui de TikTok sont des exemples assez classique, mais l'exemple suivant m'a plutôt surpris lors de mes recherches. C'est bien Ikea qui a annoncé son algorithme de recommandation hybride à l'occasion de la conférence RecSys en 2019.



Cet algorithme utilise 3 méthodes :

- Méthode du plus proche voisin (CNN pré-entraîné)
- Méthode du plus proche voisin (réseau neuronal siamois)
- Matrice de Gam

L'idée de cet algorithme est assez simple. Grouper ensemble les objets dont le style se correspond pour déterminer quels sont les produits similaires, et les produits complémentaires les uns aux autres pour ensuite les pousser aux utilisateurs.



Figure 8 - Photo de la présentation de l'algorithme de recommandation de IKEA

Source : [ARD](#)

L'algorithme de recommandation d'IKEA joue un rôle essentiel dans l'amélioration de l'expérience client de diverses manières.

Tout d'abord du côté production, il guide le développement de nouveaux produits en analysant les retours et en identifiant les tendances actuelles. Il influence directement la création de produits pour répondre aux souhaits des clients. C'est un réel assistant pour IKEA, à la fois pour créer une image de leurs consommateurs et les styles les plus appréciés. Mais également pour permettre aux chaînes de production de se concentrer sur les top ventes et de par conséquent améliorer la gestion de production et de stock. Avec ces données, on peut donc définir quels sont les modèles à moins produire pour éviter de les stocker aux détriments d'autres produit et donc de diminuer d'une certaine manière les coûts.

Et d'un autre côté, cet algorithme peut aider l'utilisateur à rester dans un style cohérent au moment de l'achat en lui recommandant des articles complémentaires qui pourrait habiller son intérieur. Des recommandations qui peuvent également pousser à l'achat. Ou au contraire élargir les propositions. Par exemple, je consulte un bureau :

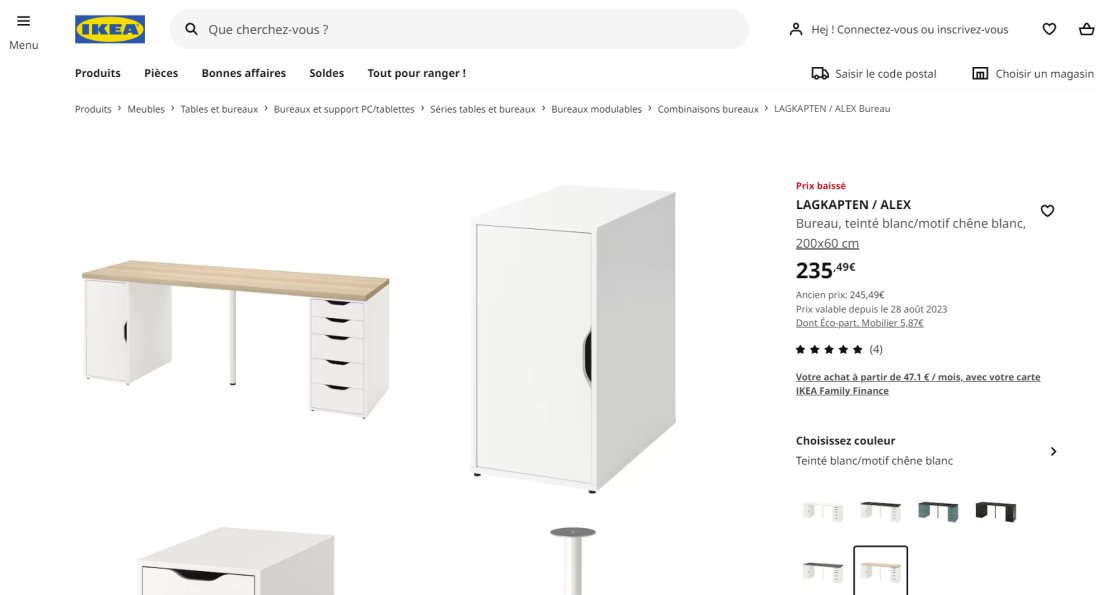


Figure 9 - Page produit d'un bureau d'IKEA

L'algorithme me propose d'autres bureau dans le même style ainsi que des articles qui pourrait être cohérent en complément de l'achat de mon bureau :

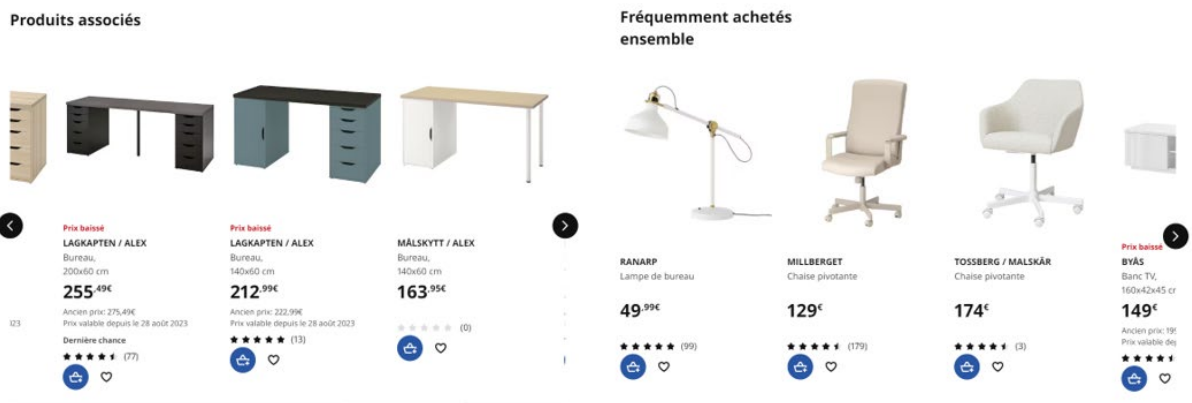


Figure 10 - Recommandation découlant de la page produit du bureau

Le cas de TikTok

Le dernier exemple est l'application la plus à la mode en ce moment : TikTok. Il se découpera en 2 parties. Premièrement on va revenir un peu sur comment ByteDance a créé une des applications les plus addictives jamais créée. Puis dans un second temps étudié comment les algorithmes de recommandation de TikTok l'ont rendu aussi populaire et aussi addictive.

Découvrons l'histoire de TikTok et de sa société mère, ByteDance. Fondée en 2012 par Zhang Yiming un ingénieur spécialisé dans les algorithmes de recommandation. Mais pas n'importe quel algorithme, des algorithmes addictifs, ByteDance a débuté avec Toutiao, une application de recommandation d'actualités alimentée par des algorithmes de recommandation. Toutiao vous propose les actualités que vous voulez avant même que vous le sachiez. Puis, en réponse à l'émergence de plateformes de partage de micro-contenu telles que Vine, ByteDance a donné vie à Douyin en septembre 2016 (version chinoise de TikTok), remportant un succès immédiat auprès des jeunes en Chine. En même temps, du côté du reste du monde, Musical.ly, une application basée sur le partage de micro-contenu se développe en Europe et en Amérique du Nord, ByteDance décide donc de lancer TikTok, une application basée sur Douyin pour se développer hors de Chine.

En 2017, ByteDance a acquis Musical.ly. Puis, en 2018, l'entreprise vient fusionner Musical.ly et TikTok dans l'objectif de conquérir le monde. C'est une fusion assez bénéfique pour l'entreprise car en prenant en compte l'écrasante domination sur les algorithmes de recommandations addictifs de TikTok et la base d'utilisateurs de Musical.ly cette nouvelle version de TikTok n'a pas d'équivalent.

Bien que ByteDance ait fait l'objet de critiques pour sa politique de censure en Chine avec Douyin. Mais ce n'est pas pour autant que TikTok est parfait et ne fait pas polémique. En 2021, le Gouvernement Chinois se mêle des affaires de ByteDance et rachète une part de l'entreprise (1%) ce qui permet à Pékin de prendre un siège au conseil d'administration ce qui pose de nombreux problèmes, notamment sur le déplacement de données personnelles vers la Chine, ... Mais on aura le temps d'y revenir.



Saviez-vous que le nom **TikTok** viendrait du tic-toc ou tic-tac de l'horloge car selon l'entreprise « chaque seconde compte ». Assez ironique venant d'une entreprise où leurs chiffres sont basés sur votre attention.

L'algorithme de TikTok est un système qui génère des recommandations personnalisées de contenus aux utilisateurs. Il se base sur une analyse précise des centres d'intérêt de chaque utilisateur. L'objectif de l'application est de faire scroller l'utilisateur sans qu'il ne se lasse.

Chaque action réalisée sur la plateforme est prise en compte pour personnaliser l'expérience utilisateur. Ces actions aident l'algorithme à devenir plus performant et à mieux comprendre les goûts de chacun. Selon les préférences des utilisateurs, TikTok propose des recommandations de contenus dans le flux #PourToi, dès l'ouverture de l'application. Cette page est alimentée par un système de recommandation qui offre à chaque utilisateur du contenu qui peut l'intéresser. Il faut également savoir que la plateforme privilégie la pertinence des vidéos et ne tient pas compte du nombre d'abonnés du créateur. Ainsi, même si une marque possède peu, voire pas du tout d'abonnés, elle peut rapidement percer sur TikTok en étant régulière dans son contenu.

Lorsqu'un créateur publie une nouvelle vidéo, l'algorithme va parcourir différentes étapes avant de la rendre virale. Il va tester cette vidéo sur une petite quantité d'utilisateur puis si elle passe le premier test, il va petit à petit augmenter le nombre de personnes à qui le contenu sera suggéré. C'est ce même système de test qui est utilisé sur la plus part des plateformes de partage de vidéo (Youtube, ...)

Pour quantifier la performance d'un contenu, TikTok s'appuie sur un barème de notation sur 30 points.

- Le taux de visionnage = 10 points
- Le taux d'achèvement (est-ce que l'utilisateur a visionné la vidéo entièrement ?) = 8 points
- Le partage = 6 points
- Les commentaires = 4 points
- Les likes = 2 points

Bien sûr tous les autres formes de feedbacks sont pris en compte.

Ces éléments sont utilisés pour évaluer la qualité et la pertinence d'une vidéo pour les utilisateurs. Les vidéos qui obtiennent un score élevé ont plus de chances d'être recommandées à d'autres utilisateurs. Si le score global de votre contenu est supérieur à la moyenne, l'algorithme TikTok le poussera à un public plus large, partageant les mêmes centres d'intérêts. En revanche, si le score est inférieur à la moyenne, votre vidéo risque de disparaître dans les abysses. Si votre contenu ne devient pas viral dès les premiers jours, ne baissez pas les bras. Le cycle ne se termine pas, il s'interrompt simplement. Cela signifie que l'algorithme pourrait repasser sur votre contenu une fois que vous commencerez à publier plus fréquemment sur la plateforme. On peut le faire sous forme



Figure 11 - Illustration du cycle d'évaluation d'un contenu sur TikTok

En ce qui concerne l'addiction à TikTok, une étude menée aux États-Unis entre avril et mai 2023 a révélé que plus de 73% des répondants ont déclaré penser que TikTok était addictif. De plus, environ 27% des répondants ont déclaré ressentir des effets négatifs sur leur santé mentale en raison de l'utilisation de l'application. Une autre étude a identifié certains comportements liés à la dépendance à TikTok, lorsque les utilisateurs développent une dépendance à l'application de vidéos virales. Bien que la majorité des utilisateurs (68,2%) soient classés comme ayant "aucun risque" de dépendance à

TikTok, 25,4% d'entre eux sont considérés comme étant à "faible risque", et 6,4% sont classés comme étant "à risque".

Il est vrai que l'interface utilisateur en plein écran et la présentation directe du contenu lors de l'ouverture de l'application peuvent contribuer à l'addiction à TikTok. De plus, le fait de limiter le nombre de choix conscients faits par l'utilisateur peut également contribuer à cette addiction, car cela peut donner l'impression que l'utilisateur est "piégé" dans l'application.



Figure 12 - Intégration de l'interface utilisateur de TikTok

On a donc vu que les algorithmes de recommandation sont de plus en plus utilisés dans diverses applications pour améliorer l'expérience utilisateur. Les exemples mentionnés, Spotify, Ikea et TikTok, sont des exemples de la façon dont les algorithmes de recommandation peuvent être utilisés pour en théorie améliorer l'expérience utilisateur.

Spotify utilise un algorithme de recommandation pour recommander de la musique personnalisée à ses utilisateurs en fonction de leurs goûts musicaux. Ikea utilise un algorithme de recommandation pour recommander des produits en fonction des préférences de ses clients. TikTok utilise un algorithme de recommandation pour recommander des vidéos personnalisées à ses utilisateurs en fonction de leurs centres d'intérêt, les rendant addict.

Ces algorithmes de recommandation sont conçus pour améliorer l'expérience utilisateur en offrant des recommandations personnalisées qui correspondent aux goûts et aux préférences de chaque utilisateur. Cependant, il est important de noter que ces algorithmes peuvent également être addictifs et avoir des effets négatifs sur la santé mentale des utilisateurs. Il est donc important de les utiliser avec prudence et de prendre des pauses régulières pour éviter la dépendance. Mais l'enjeu des données est également très important. C'est le sujet du prochain chapitre.

CHAPITRE 4

Données et éthique : les enjeux des algorithmes de recommandation

Dans ce dernier chapitre avant de conclure, je vais vous parler d'un aspect fondamental : les données. Au cœur des algorithmes de recommandation, les données jouent un rôle essentiel. Ce chapitre se penche sur la nécessité de ces données pour alimenter ces algorithmes, mettant en lumière les enjeux associés à leur utilisation.

Comme vous l'avez vu dans les divers chapitres précédents, les données sont vraiment au cœur des algorithmes de recommandations car sans données, il n'existe aucun moyen de personnaliser l'expérience. Mais au-delà de personnaliser l'expérience de l'utilisateur sur lequel les données sont collectées, mais elles permettent également de personnaliser le contenu d'autres utilisateurs avec des mécanismes comme le filtrage collaboratif (CF). L'exemple de TikTok montre également que la collecte de données peut causer des problèmes géopolitique, à cause du transfert de données entre pays (ici la Chine). Pour tenter de réguler la collecte de données personnelles, certains dispositifs sont mis en œuvre :

- Le **Règlement Général de Protection des Données (RGPD)** en Europe, entrée en vigueur le 25 mai 2018. Elle vise à réglementer la collecte et le traitement des données personnelles, elle s'applique à toutes les entreprises qui collectent et traitent des données personnelles de citoyens européens.
- La **California Consumer Privacy Act (CCPA)** est une loi californienne qui régleme la collecte et l'utilisation des données personnelles des consommateurs californiens.
- La Loi sur la **Protection des Renseignements Personnels et les Documents Electroniques (LPRPDE)** au Canada.
- Loi sur la **protection des données personnelles (PDPA)** à Singapour.

Il existe une carte sur le site de la CNIL qui montre l'état de protection des données dans le monde :

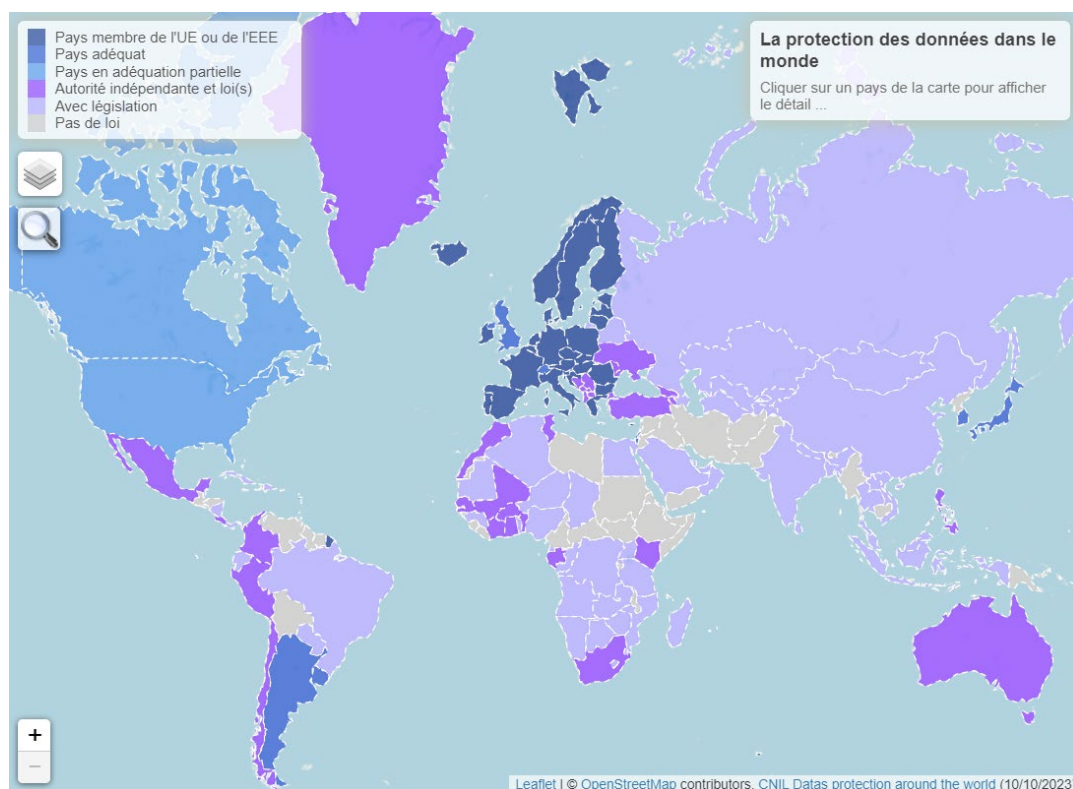


Figure 13- Carte de protection des données dans le monde (Source CNIL)

Depuis quelques années, la plus part des applications comme Instagram ou TikTok proposent de télécharger ses données pour pouvoir se rendre compte de l'ampleur de cette collecte. Pour vous sensibiliser à la quantité de données que ces services collectent j'ai décidé de télécharger mes données. J'ai pu télécharger mes données sur les plateformes :

- Instagram
- TikTok

Pour commencer avec mes données d'Instagram, vous pouvez donc les télécharger depuis le lien suivant : https://accountscenter.instagram.com/info_and_permissions/dyi/

On peut donc lister différents types d'informations collectées :

- L'ensemble de mes conversations (boite de réceptions et demandes)
- La liste de mes sujets d'intérêt
- Les interactions que j'ai pu avoir avec des stories (likes, quizz, questions, ...)
- Les contenus enregistré
- Mes informations personnel (compte, profil, ...)
- Mes informations de monétisation
- Tous les contenus liké et commentaires likés
- Les informations des appareils sur lesquels j'ai été connecté (modèle, OS, localisation, ...)
- Tous les contenus que j'ai publié (stories, réels, publications, photos de profil, ...) on trouve à la fois l'image publiée et les données associées (métadonnées, description, ...)
- Mes contacts synchronisé avec l'application
- Tous les commentaires que j'ai pu rédiger
- Les données associé à mon profil sur les publicités
- Mon historique de recherches (même si supprimé)
- Les comptes suivies et mes followers
- ...

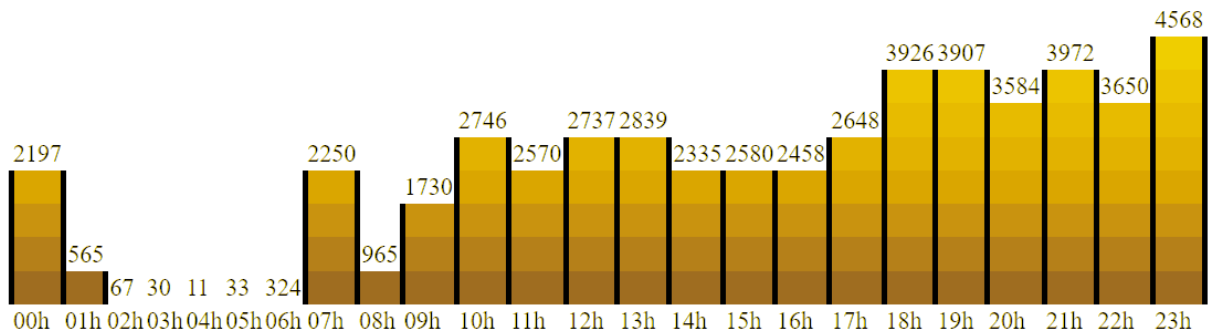
Personnellement, j'ai pu traiter quelques une de ces données avec un code python :

J'ai pu extraire quelques chiffres et graphiques de mon utilisation :

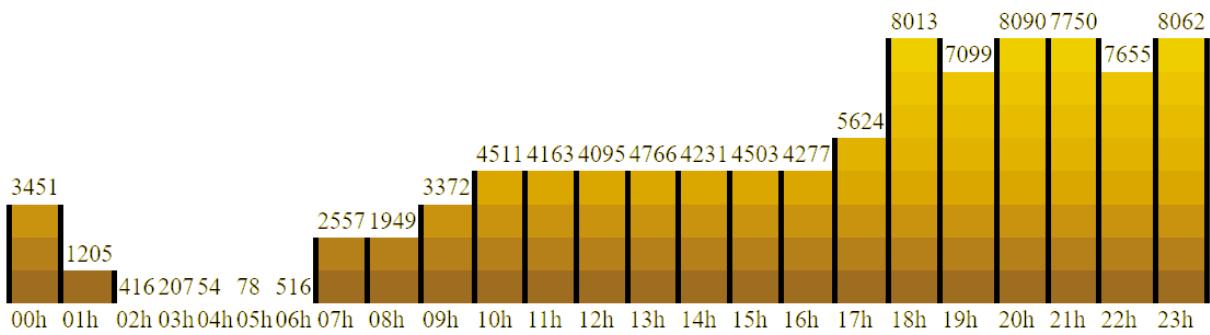
- 52 692 messages envoyés
- 96 644 messages reçus
- 308 conversations
- Mon jour le plus actif (16 mars 2020) avec 383 messages envoyés
- Ma période la plus active de la journée (23h) avec 8% des messages

J'ai également pu extraire des données pour les afficher sous formes de graphiques :

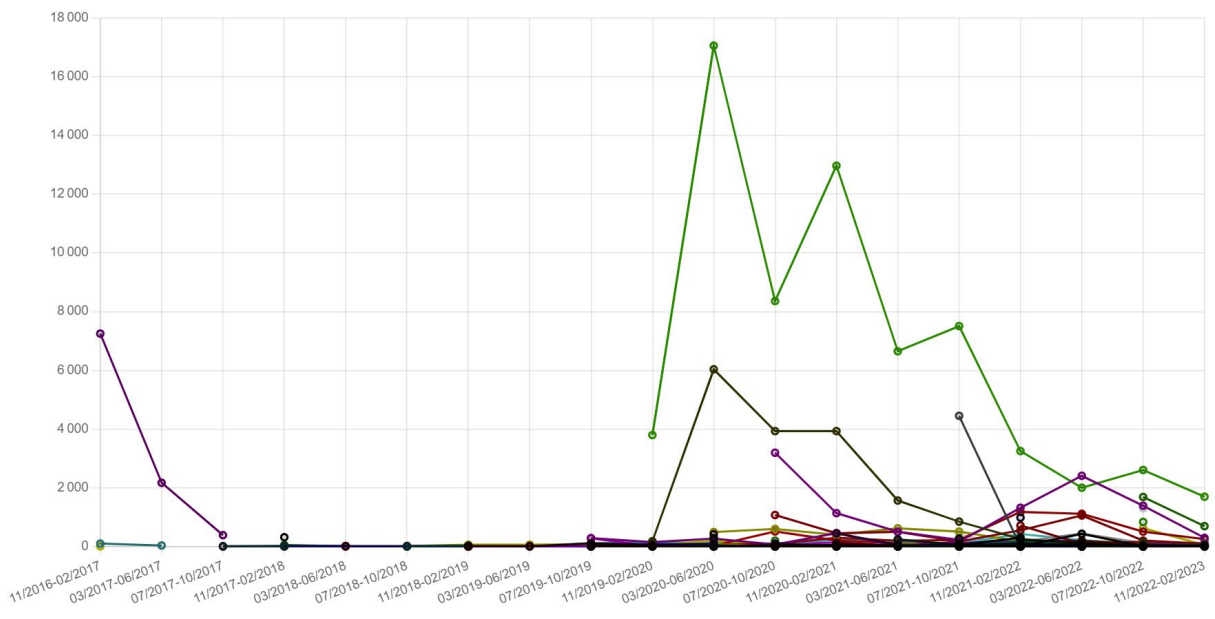
Nombre de messages envoyés par heure



Nombre de messages reçus par heure



Evolution du nombre de messages au cours du temps



Et c'est une partie minime des données téléchargées.

Pour ce qui est de mes données TikTok, vous pouvez également les télécharger depuis le lien suivant : <https://www.tiktok.com/setting/download-your-data>

On peut donc lister différents types d'informations collectées :

- Activité du compte
 - Contenu ajouté en favoris
 - Effets
 - Hashtags
 - Musiques
 - Vidéos
 - Liste des comptes suivis
 - Liste des followers
 - Liste des comptes bloqués
- Interactions
 - Liste des hashtags utilisés
 - Liste des contenus likés
- La localisation la plus récente (date, données GPS, région de la localisation)
- Différents historique
 - Historique des connexions (date, adresse IP, model de l'appareil, OS de l'appareil, type de réseau, fournisseur d'accès)
 - Historique d'achat
 - Historique des recherches (même supprimé encore une fois)
 - Historique des partage
- Une liste de statuts (résolution d'affichage, version de l'application, IDFA, GAID, Android ID, IDFV, Web ID)
- L'ensemble des vidéos consultés (date et lien)
- Information publicitaire
 - Les centres d'intérêts
- Les paramètres utilisateurs de l'application
 - Autorisation du téléchargement
 - Langage
 - préférences de contenus
 - le filtre de commentaires
 - le statuts du compte
 - Etat des notifications (ordinateurs, nouveau contenu, nouveau follower, ...)
 - Paramètres de relation aux autres utilisateurs
 - Qui peut faire des 'duo' avec moi
 - Qui peut commenter mes publications
 - Qui peut me parler par messages
 - ...
- L'ensembles des commentaires (date, commentaire, contenu lié)
- L'ensemble des conversations (date, message, interlocuteurs, ...)
- Information de profil
 - numéro de téléphone

- email
- nom
- prénom
- adresse (rue, code postal, pays, ...)
- description de profil
- data d'anniversaire
- like reçu
- photo de profil
- ...
- Informations lié au live TikTok
 - Paramètres
 - Historique des live regardé (commentaire, questions, watch-time, lien)
- Mes contenus publiés (date, lien, likes, text alternatif, ...)
- ...

Et cette liste n'est pas effective, le compte duquel sont tiré ces données est un compte à usage professionnel, d'un utilisateur très peu consommateur du service TikTok.

La course aux données est devenue un enjeu majeur pour les entreprises, car les algorithmes ont besoin de ces données pour fonctionner efficacement. Cependant, les utilisateurs ont également besoin de se savoir protégés sur le plan numérique. Les géants de la technologie tels que Google, Meta et d'autres sont souvent accusés de survoler les lois pour obtenir ces données.

Meta, l'entreprise propriétaire de Facebook, a été sanctionnée par la CNIL irlandaise pour avoir enfreint les règles européennes sur la protection des données. Meta a été condamné à une amende record de 1,2 milliard d'euros pour avoir « continué de transférer des données personnelles » d'utilisateurs de l'Espace économique européen (EEA) vers les États-Unis en violation des règles européennes en la matière.

Cependant, il est clair que ces lois doivent être renforcées pour garantir une meilleure protection des données personnelles. Les géants de la technologie tels que Google, Meta et d'autres doivent respecter ces lois et être tenus responsables de leurs actions.

CONCLUSION

Les algorithmes de recommandation sont omniprésents dans notre vie quotidienne, de notre réveil avec notre actualité numérique jusqu'à notre moment détente sur les réseaux sociaux.

Cette veille technologique avait pour ambition de présenter la notion d'algorithme de recommandation de manière simple, de comprendre comment fonctionne cette technologie et de sensibiliser sur les différents enjeux qui gravitent dans cet univers des recommandations. En résumé, les algorithmes de recommandation dirigent nos choix en ligne, rendant notre expérience plus personnalisée. Cependant, il est crucial de s'interroger sur la quantité de données que nous partageons pour alimenter ces recommandations.

Les algorithmes de recommandation ont le potentiel de transformer la façon dont nous interagissons avec le contenu et les autres utilisateurs sur les plateformes en ligne. Ils peuvent influencer nos choix, nos opinions et nos comportements, et peuvent avoir des impacts significatifs sur notre vie privée et notre société. Il est donc important de comprendre les enjeux liés à l'utilisation de ces algorithmes et de mettre en place des mesures pour minimiser les risques.

Cette veille technologique vous invite à vous questionner : Quels sont les impacts des algorithmes de recommandation sur notre vie quotidienne ? Quelles mesures pouvons-nous prendre pour une utilisation responsable et éthique des algorithmes de recommandation ?



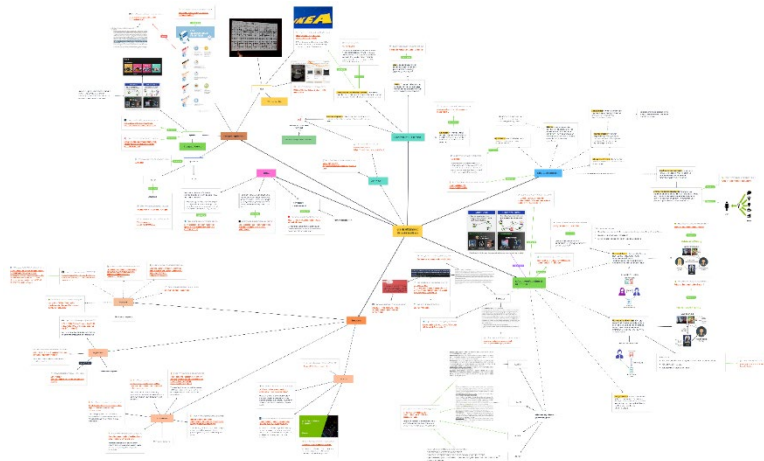
Les algorithmes de recommandation ne sont pas une fin en soi, mais un moyen de mieux comprendre les besoins et les préférences des utilisateurs. Ils peuvent nous aider à découvrir de nouveaux contenus, à élargir nos horizons, et à nous connecter avec des personnes qui partagent nos intérêts. Mais pour cela, nous devons être conscients de leur potentiel et de leurs limites, et travailler ensemble pour les améliorer.

BILAN

Transposition de cette technologie en entreprise, ouverture, plan de diffusion et prise de recul

Retour sus mes pistes de recherches

Malgré mon intérêt déjà présent sur la thématique des algorithmes de recommandation j'ai tout de même du compléter mes recherches notamment sur leurs fonctionnement interne. Sur les formules mathématiques qui se cachent derrière ces recommandations. Pour ce faire j'ai construit une carte mentale assez complète comprenant tous les axes que je souhaitais aborder. Cette carte mentale m'a permis d'organiser mes différentes lectures, et informations collectés ainsi que leurs ressources mais également de faire des liens entre les items. J'ai également dû abandonner certaines pistes par manque de pertinence ou de sources.



Projection sur l'application des algorithmes de recommandations en entreprise

Je ne suis pas sûr que ces algorithmes de recommandation soit voué à être implémenté au sein de l'entreprise. Je pense que ces systèmes de recommandation vienne en complément d'une offre ou d'un service fournit par l'entreprise. Un système de recommandation peut être intéressant à mettre en place si l'entreprise propose une large gamme de produit. Si l'entreprise propose 3 offres, je ne suis pas sûr qu'il y ai une grande plus-value, un simple tableau comparatif serait plus évident pour l'acheteur final.

Ouverture sur les problématiques liées aux algorithmes de recommandations

Comme exposé dans les chapitre 3 et 4, les algorithmes de recommandation pose de nombreux problème et notamment sur la gestion des données. Mais ce problème n'est pas uniquement pour ces algorithmes, c'est un problème présent depuis l'arrivée du numérique. Par le passé, nous avons déjà été confronté à de gros problème de collecte de données, une des affaires les plus connu et celle des révélations faites par Snowden ou plus récemment avec l'émergence des LLM (Large Langage Model), des modèles souvent entraîné sur des quantités astronomique de données. L'enjeux des données est un des axes d'amélioration important de ces technologie. C'est ce même besoin en données qui pose des problèmes de démarrage à froid ou de surpersonnalisation. Tout tourne autour des données. Un autre problème présent avec ces algorithmes et qui a été mis en avant dans la présentation de TikTok, c'est l'addiction que peut générer ces algorithmes. En proposant du contenu toujours plus proche de nos attentes, ces géants ont réellement mis un Bouton a dopamine dans nos poches. Cette addiction a la dopamine n'est pas nouveau. « En 2016 un neurobiologiste (Christian Lüscher) met des souris dans une cage avec un bouton. Quand elles appuient sur ce bouton un capteur stimule leur cerveau pour qu'il libère de la dopamine, la molécule associée au plaisir ; chaque fois qu'une souris appuie sur le

bouton elle reçoit un shot de joie instantané. La dopamine permet de ressentir un plaisir, le cerveau aime tellement cette substance que, s'il comprend que manger du chocolat égale « plaisirs », alors il peut en devenir dépendant. Rapidement les souris deviennent addict à ce bouton, elles veulent recevoir de plus en plus de shot de dopamine au point qu'elles ne pensent plus à dormir, à manger, ni même à se reproduire. Si les scientifiques ne les avaient pas sorties de leur cage les souris seraient mortes heureuses ! » Chasseurs d'Actu (Cf. Sources)

Les algorithmes de recommandations pose donc de nombreux problèmes, mais la question est : Est-ce que la personnalisation de l'expérience utilisateur en vaud le coup. Pour eux la réponse est oui, car leurs économie est basé sur votre attention. Et vous, quel est votre avis.

Sources sur l'étude de Christian Lüscher :

- European Scientist - Des chercheurs suisses mettent en évidence le circuit cérébral de l'addiction : [Lien](#)
- Université de Genève - Le circuit cérébral qui détient la clé de l'addiction : [Lien](#)
- Chasseurs d'Actu – TikTok, la drogue 2.0 ? [Lien](#)

Sources sur l'affaire Snowden :

- Le Monde - Ce que les « révélations Snowden » ont changé depuis 2013 : [Lien](#)
- Wikipédia – Révélation d'Edouard Snowden : [Lien](#)
- Lumni - Révélations d'Edward Snowden sur le programme de surveillance de la NSA : [Lien](#)
- Film – Snowden (2016) : [Lien](#)

Plan de diffusion de mon travail

- **Outils de diffusion**
Partage sur LinkedIn et mon site web en version intégrale et extrait spécifique.
- **Objectif de la communication**
Faire découvrir mon profil et sensibilisé les lecteurs sur leurs données et ce qui se cache derrière les recommandations en ligne.
- **Public cible**
Etudiant en informatique, professionnel du domaine ou tout autre personne intéressée par le sujet.
- **Objectifs**
Augmenter la visibilité de votre travail, attirer de nouveaux visiteurs sur votre site web, et établir votre expertise dans le domaine des algorithmes de recommandation.
- **Indicateurs de performance (KPI)**
Nombre de vues, partages, commentaires, et mentions "J'aime" sur LinkedIn ; nombre de visites et de téléchargements du rapport sur votre site web.

Réflexion personnelle

Cette étude m'a permis de mieux comprendre le réel fonctionnement des algorithmes de recommandation mais également d'approfondir mes connaissances en Big Data. Je suis très satisfait de la qualité de mon rapport et je pense avoir la meilleure version de mon rapport possible.

BIBLIOGRAPHIE

une plongée dans les détails de mes sources

- Wikipédia. (2023, 29 décembre). Système de recommandation. [Lien](#)
- Algorithme. (s. d.). CNIL. [Lien](#)
- IA et algorithmes de recommandation web. (s. d.). [Lien](#)
- Intelligence artificielle, de quoi parle-t-on ? (s. d.). CNIL. [Lien](#)
- Wikipédia. (2023, 7 mars). Filtrage collaboratif. [Lien](#)
- What Content-Based Filtering is and Why You Should Use It. (2021, 6 avril). upwork. [Lien](#)
- What is a recommendation system ? (s. d.). NVIDIA Data Science Glossary. [Lien](#)
- Carro, B. (2022, 10 janvier). Quels sont les principaux algorithmes de recommandation ? Mediego. [Lien](#)
- Des algorithmes et des hommes : recommandation personnalisée des contenus. (s. d.). [Lien](#)
- Les évolutions du Web : le Web 1.0, le Web 2.0, le Web 3.0 et le Web 4.0. (2021, 11 octobre). Click2Cell - Agence web & digitale à Clermont-Ferrand et en Ile-de-France. [Lien](#)
- YouTube. (s. d.). Vulgarisation sur l'IA. [Lien](#)
- Etude des systèmes de recommandations et mise en pratique des algorithmes. (2020). Mathéo, 15. [Lien](#)
- Talbi, I. (2020, 10 janvier). Les algorithmes de recommandations mettent en danger votre liberté. Forbes France. [Lien](#)
- Wikipédia. (2024, 5 janvier). Netflix. [Lien](#)
- Wikipédia. (2023, 27 décembre). Amazon. [Lien](#)
- Rédaction, L. (s. d.). Amazon : Qu'est-ce que c'est ? Futura. [Lien](#)
- Wikipédia. (2023b, 5 décembre). Spotify. [Lien](#)
- Wikipédia. (2023c, décembre 21). Flickr. [Lien](#)
- Wikipédia. (2023c, décembre 15). YouTube. [Lien](#)
- Wikipédia. (2021, 27 septembre). Friendster. [Lien](#)
- Wikipédia. (2022, 2 novembre). Myspace. [Lien](#)
- Wikipédia. (2023b, décembre 3). Usenet. [Lien](#)
- Larousse, É. (s. d.-a). Définitions : Éthique - Dictionnaire de français Larousse. [Lien](#)
- Wikipédia. (2023b, juin 14). Interactivité. [Lien](#)
- Rédaction, L. (2009, 23 septembre). Netflix met fin à son concours doté d'un million de dollars. [Lien](#)
- DRUMOND. G, COUNTANT. A, & MILLERAND. F. (2018, 15 novembre). La production de l'utilisateur par les algorithmes de Netflix. Les Enjeux de l'information et de la communication. [Lien](#)
- Schwab, P. (2020, 27 mai). Comment hacker les algorithmes de recommandation ? Conseils en marketing. [Lien](#)
- Algorithmes de recommandation : la magie derrière vos choix | Enoki Studio. (s. d.). enoki-studio.com. [Lien](#)
- Les sites web utilisant des algorithmes de recommandation pour améliorer l'expérience utilisateur. (2023, 23 octobre). Webady. [Lien](#)
- Schwab, P. (2023, 12 juillet). Analyse des algorithmes de recommandation : l'intérêt des méthodes mixtes. Conseils en marketing. [Lien](#)
- Contributeurs aux projets Wikimedia. (2023e, décembre 5). Spotify. [Lien](#)
- Wikipedia (2023, 3 février). The Echo Nest. [Lien](#)
- Les algorithmes de recommandation de Spotify : de la découverte aux bulles de filtres culturelles. (2019, 13 mars). LES MONDES NUMERIQUES. [Lien](#)
- Duvall, W. (2023, août 4). 4 things Every artist needs to know about Spotify. recommendation algorithm. Playlist Pousser. [Lien](#)

- Eggleston, B. (2023, 16 mai). The Spotify algorithm : How to get on release radar. Playlist Poussier. [Lien](#)
- Pete. (2024, 8 janvier). Challenges faced by modern record labels. LabelGrid. [Lien](#)
- TechCrunch is part of the Yahoo family of brands. (2017, 18 mars). [Lien](#)
- FRENCHWEB.FR. (2014, 6 mars). Musique : Spotify s&rsquo ; offre le moteur « big data » The Echo Nest. FRENCHWEB.FR. [Lien](#)
- Challenges.fr. (2014, 6 mars). Spotify achète la plateforme musicale intelligente The Echo Nest. Challenges. [Lien](#)
- Wikipédia (2023h, décembre 21). The playlist. [Lien](#)
- Susic, P. (2023, 30 septembre). 20+ Music Genre Statistics : Most popular music genres (2024). [Lien](#)
- Shewale, R. (2023, 28 décembre). Spotify Stats for 2024. [Lien](#)
- Amy. (2021, 16 mai). Retours de Copenhague 2 : Algorithmes pour la déco. Acheter, Rénover, Décorer. [Lien](#)
- Schwab, P. (2019, 19 novembre). IKEA présente un système de recommandation hybride. Conseils en marketing. [Lien](#)
- Defend Intelligence. (2022, 16 novembre). Le plus gros problème avec TikTok. [Vidéo]. YouTube. [Lien](#)
- Obligations en matière de protection des données personnelles. (2022, 1 avril). Entreprendre.Service-Public.fr. [Lien](#)
- Garantir la sécurité des données. (s. d.). CNIL. [Lien](#)
- Les principes fondamentaux de la protection des données personnelles - Inserm Pro. (2021, 9 juillet). Inserm pro. [Lien](#)
- Solutions de sécurité et de protection des données | IBM. (s. d.). [Lien](#)
- La protection des données dans le monde. (s. d.). CNIL. [Lien](#)
- En Europe et dans le monde. (s. d.). CNIL. [Lien](#)
- La protection des données dans le monde. (s. d.-b). CNIL. [Lien](#)
- Contributeurs aux projets Wikimedia. (2023i, décembre 22). Meta (entreprise). [Lien](#)
- Fabrian, M. (2023, 5 janvier). RGPD : Meta commence 2023 avec une amende de 390 millions d’euros. Les Numériques. [Lien](#)
- Latribune.fr. (s. d.). RGPD : Meta condamné à une amende de 1,2 milliard d’euros, la plus élevée jamais infligée en Europe. La Tribune. [Lien](#)
- Protection des données dans l’UE. (s. d.). Commission européenne. [Lien](#)
- Statista. (2023, 5 septembre). TikTok Mental health Effects Opinions among U.S. users 2023. [Lien](#)
- Smith, A. (2022, 5 mai). New study identifies ‘TikTok addiction’ and the symptoms that show when users have it. The Independent. [Lien](#)
- Gué, V. (2024, 8 janvier). Comment fonctionne l’algorithme TikTok en 2024 ? Hubspot. [Lien](#)
- Patard, A. (2020, 25 février). Comment fonctionne l’algorithme de TikTok. BDM. [Lien](#)
- Statista. (2023c, septembre 5). TikTok Mental health Effects opinions among U.S. users 2023, by generation. [Lien](#)
- Biggs, J. (2022, 19 mai). TikTok Addiction : Why is TikTok so addictive ? Cosmopolitan. [Lien](#)
- Léo Duff. (2023, 5 février). Comment TikTok a ba*sé le cerveau d’une génération. [Lien](#)
- Crossley, C. (2022, 21 mars). TikTok’s addictiveness is a problem — especially for young minds. GBH. [Lien](#)
- Paul, K. (2022, 30 octobre). What TikTok does to your mental health : ‘It’s embarrassing we know so little’. the Guardian. [Lien](#)

- TikTok is killing your brain, one short-form video at a time. (2022, 18 septembre). Social Media Psychology. [Lien](#)
- Davis, M. (2021, 25 octobre). TikTok is bad for your brain : constant social media streaming narrows collective attention span, adversely affects mental health. Science Times. [Lien](#)
- Collaborative Filtering Recommender Systems. (2011). GroupLens. [Lien](#)
- Short history of collaborative filtering - information retrieval, recommendations, Tapestry, Amazon, Ringo, Firefly. (s. d.). [Lien](#)
- Lebrun, S. (2023, 17 mai). Comment fonctionne l& # 39 ; algorithme TikTok ? [Lien](#)
- Signoret, P. (2019, 8 novembre). Au fait, pourquoi TikTok s'appelle TikTok ? Numerama. [Lien](#)
- Wikipédia. (2024b, janvier 15). TikTok. [Lien](#)
- Home - RecSys. (2024, 10 janvier). RecSys. [Lien](#)
- Bechade, C. (2020, 20 octobre). La nouvelle guerre des étoiles : tous notés, partout, tout le temps ? Les Numériques. [Lien](#)
- Avide de notes sur ses produits, Amazon teste l& # 039 ; évaluation en un clic. (2019, 17 septembre). Clubic.com. [Lien](#)
- What Content-Based Filtering is and Why You Should Use It. (2021b, avril 6). UpWork. [Lien](#)
- Des algorithmes et des hommes : recommandation personnalisée des contenus. (s. d.-c). calameo.com. [Lien](#)
- Casalegno, F. (2022, 12 décembre). Recommender Systems — a complete guide to machine learning models. Medium. [Lien](#)
- Minds, C. (2022, 8 septembre). What are today's top recommendation engine algorithms ? Medium. [Lien](#)
- Bastin, & Jérémy. (2021). Etude des systèmes de recommandations et mise en pratique des algorithmes. Liège Université. [Lien](#)
- Chatelain, Y. (2022, 2 novembre). Algorithmes de recommandation, outils de radicalisation. Contrepoints. [Lien](#)
- ByteDance - Inspire creativity, enrich life. (s. d.). [Lien](#)
- Joanna. (2023, 1 décembre). Les systèmes de recommandation favorisent-ils la diversité ? - Interstices. Interstices. [Lien](#)
- Algorithme. (s. d.-b). CNIL. [Lien](#)
- Short history of collaborative filtering - information retrieval, recommendations, Tapestry, Amazon, Ringo, Firefly. (s. d.-b). [Lien](#)
- Fabien Olicard La deuxième chaîne. (2023, 4 juin). Pourquoi TikTok a (déjà) détruit votre futur (annonce de mon nouveau livre) [Vidéo]. YouTube. [Lien](#)



ROBIN NOIRET

Etudiant CESI

Site : robinnoiret.fr

LinkedIn : Robin Noiret

Instagram : @inspire_by_robin

