

IA, La machine à livrer

Pedro Garcia, Kais Riahi, Marlan Sadvakas, Daniel Lipp

Etudiant-e-s en ingénierie des médias, 1^{ère} année, HEIG-VD

De nos jours, les entreprises du domaine des médias sont obligées d'utiliser l'intelligence artificielle (IA) au service des consommateurs de contenu afin de fournir efficacement leurs produits sur le marché. Sur Netflix, par exemple, plus de 80 % des contenus consommés sont découverts via le système de recommandation de la plateforme, celui-ci étant entièrement conçu à l'aide d'algorithmes dits intelligents, capables d'analyser, traiter et fournir du contenu ciblé. Ceux-ci appliquent la notion de « Machine Learning ». Ce mécanisme est capable de « deviner », avec précision, les envies d'un utilisateur en observant son comportement et une série d'autres points qui seront traités dans cet article. Le pouvoir de choisir, perdra-t-il sa valeur un jour ?

I. INTRODUCTION

La population fait face à un nombre croissant de choix dans chaque aspect de leur vie. Lors du développement des technologies appliquées aux médias, soit des algorithmes capables d'améliorer la modélisation, la promotion et la diffusion du contenu; nous avons de plus en plus la possibilité de choisir quoi, quand, où et comment en bénéficier. Des vidéos, de la musique, des livres, des hôtels, des restaurants, etc., les possibilités ne manquent pas.. Par conséquent, la concurrence entre les entreprises qui fournissent ces services augmente. Les obligeant ainsi à investir dans des technologies plus sophistiquées afin de se démarquer. La rapidité d'accès, la qualité du contenu et, surtout, les meilleures recommandations sont ainsi des points à prendre en compte par les fournisseurs de contenu média.

II. TEXTE PRINCIPAL

A. LES MÉDIAS ET LE E-COMMERCE

Les médias sont des outils qui permettent de diffuser une information. Ils sont les intermédiaires entre l'émetteur et le récepteur. Aujourd'hui, le récepteur est aussi appelé « utilisateur ». C'est un terme qui désigne une personne utilisant un système informatisé pour accéder à l'information. Dans le contexte de ce travail, nous traiterons de l'utilisateur humain qui n'a aucune compétence en informatique et qui utilise ces systèmes dans le cadre de ses loisirs (e.g. regarder un film sur Netflix, voir une émission sur YouTube, écouter un album sur Spotify, etc.). Nous intégrerons aussi des exemples de fonctionnement de quelques mécanismes de recommandation de la plateforme Netflix.

Dans une diffusion commerciale d'un média, il faut prendre en compte le type d'utilisateur décrit ci-dessus (l'utilisateur humain). Ce dernier adopte des comportements, qui au sens commercial, sont d'une grande importance à observer afin de pouvoir lui diffuser du contenu ciblé. Les systèmes de recommandation sont un moyen de mettre en avant la qualité de la diffusion et le choix de l'utilisateur, en comparant certains de ses comportements avec le profil d'un utilisateur « type ». Ces systèmes cherchent à prédire quel avis donnerait un consommateur sur le contenu.

Prenons l'exemple de l'entreprise américaine Netflix, qui propose des films et séries télévisées en flux continu sur Internet et qui est l'une des principales bénéficiaires de ce mécanisme. En quelques années, celle-ci est devenue la plus grande plateforme de vidéo en streaming par abonnement. Elle ne cesse d'investir dans des productions originales pour se démarquer des concurrents qui sont des chaînes câblées. Ces productions ont été imaginées depuis les préférences des utilisateurs de la plateforme.

Cette dernière a aussi popularisé la pratique du « binge-watching », qui consiste à regarder en continu plusieurs épisodes d'une série dans un court laps de temps. Pour en arriver là, la société a collecté un nombre grandissant de données afin de proposer les suggestions les plus personnalisées possible aux utilisateurs. Plus de la moitié des contenus lus sur Netflix découlent de leur système de recommandations. Ainsi, l'utilisateur est devenu plutôt passif et les choix portés sur le prochain film ou la prochaine série qui sera jouée sont en grande partie déterminés par un algorithme.

B. LES ALGORITHMES DE RECOMMANDATION DE NETFLIX

Un des nombreux algorithmes qu'utilise l'entreprise est conçu avec la notion de l'apprentissage automatique (en anglais *machine learning*) ou apprentissage statistique. C'est un champ d'étude de l'intelligence artificielle qui concerne la conception, l'analyse, le développement et l'implémentation de méthodes permettant à une machine de remplir des tâches d'une manière plus optimale. Sur Netflix, cette méthode dresse une analyse comportementale.

Celle-ci se base sur les points suivants : la navigation dans le site ; le temps de lecture pour une série afin de déterminer le niveau d'appréciation de l'utilisateur ; les séries auxquelles l'utilisateur s'est abonné et au bout de combien de temps ; les genres de séries regardées en tentant d'y trouver une corrélation ; l'heure et la période de l'année à laquelle un contenu a été regardé.

Quand on s'intéresse au fonctionnement des mécanismes de recommandation chez *Netflix* on constate que ce dernier est d'une grande complexité. Selon un article publié en décembre 2015 par le responsable de produit de cette société, *Neil Hunt*, « le temps de séduction est compté », c'est-à-dire le temps que les utilisateurs mettent à choisir un film. L'article souligne que « la fidélité de l'abonné est l'obsession des ingénieurs de l'entreprise. Ainsi le système de recommandation est entièrement dédié à ce seul objectif ».

Netflix connaît tout à propos de ses utilisateurs. Ce qu'ils regardent, quand, la manière dont ils le font, l'intensité et surtout la façon dont ils trouvent leur contenu. Autant de données qui ont permis d'améliorer le système de recommandation de l'entreprise qui semble ne plus du tout se reposer sur les films les mieux notés, mais plutôt sur la quantité de visionnage que produit chaque vidéo du catalogue.

En effet, au début de son histoire, « l'algorithme de recommandation » de *Netflix* se reposait essentiellement sur les notes que les utilisateurs donnaient aux films. Désormais, plutôt que de montrer une évaluation entre 1 et 5 étoiles, l'entreprise propose un pourcentage représentant la probabilité que le consommateur apprécie le contenu qu'elle suggère. Dès lors, nous retrouvons un processus d'analyse d'informations qui a pour but fournir ce que le client désire d'une manière plus optimale.

C. LES MÉCANISMES DE RECOMMANDATION

La combinaison d'un algorithme qui détermine la meilleure recommandation selon le comportement de l'utilisateur avec un algorithme de « tendances récentes » permet de prendre en compte une dimension sociale liée à l'actualité dans les systèmes de recommandation.

Le « *Top-N Video Ranker* » concerne un algorithme d'analyse développé par *Netflix* qui détermine la recommandation et personnalise la page d'accueil de chaque utilisateur selon son comportement. Il y a deux types de tendances que ce « *ranker* » identifie : (1) ceux qui se répètent chaque année, comme la hausse d'observation de films romantiques pendant le jour de la *Saint-Valentin* par exemple et (2) des événements uniques, à court terme, comme un grand ouragan qui menace une certaine zone très peuplée. Étant couvert par beaucoup de médias, ceci conduit naturellement un intérêt à des documentaires et des films d'ouragans et d'autres catastrophes naturelles [fig.1]

Un autre algorithme qui fait aussi partie des systèmes de recommandation de la société est le « *Continue watching* », qui consiste à recommander du contenu selon ce qui a déjà été visionné en prenant en compte le temps écoulé depuis le visionnage d'un épisode ou le point d'abandon des séries.

L'algorithme « *Video-Video Similarity* » (similitude vidéo à vidéo), active une recommandation depuis une des vidéos qui est en train d'être visionnée par un utilisateur. Celui-ci est un algorithme non personnalisé qui pointe vers des vidéos semblables calculées pour chaque vidéo du catalogue.

Un autre algorithme encore assure la diversité de la page d'accueil et de la recommandation. Cet algorithme de construction de page est basé sur un modèle qui utilise la technique de marketing « *A/B testing* ». Ce dernier consiste à proposer plusieurs variantes d'une même page d'accueil qui diffèrent selon un seul critère et ce, afin de déterminer la version qui donne les meilleurs résultats auprès des abonnés.

D. LE MARCHÉ DES RECOMMANDATIONS

Cependant, selon *Netflix*, il reste encore des améliorations à apporter. Pour l'instant, la personnalisation de la page d'accueil consiste en un groupe de vidéos disposées en des rangées horizontales. L'application souhaite que ses abonnés prennent conscience de la manière dont elle s'adapte à leurs préférences. Ceci promeut non seulement la confiance en son système, mais encourage les membres à donner le retour d'information qui aboutira à des meilleures recommandations. Une autre façon de promouvoir la confiance avec ce mécanisme de personnalisation est de fournir des explications pour lesquelles nous décidons de recommander un film donné.

Afin d'affiner leur système de recommandation, *Netflix* a payé des gens pour observer des films et pour ensuite les étiqueter avec toutes sortes de métadonnées. Ce processus est si sophistiqué et précis que les taggeurs reçoivent un document de formation de 36 pages leur enseignant comment noter le contenu des films. Ils capturent ainsi des douzaines d'attributs de film différents. Ils notent même le statut moral des caractères. L'objectif principal de l'entreprise est rappelons-le, d'avoir le plus d'abonnés possible et de les conserver.

Pour promouvoir un film ou une série, *Netflix* utilise une petite astuce : changer la vignette de présentation de ceux-ci. Cela peut paraître peu pertinent mais étant donné que l'abonné ne passe en moyenne que 1,8 secondes pour déterminer quel film il regardera, l'image qui attirera son attention compte pour 82 % de cet espace-temps.

Tous ces algorithmes et mécanismes d'analyse décrits ci-dessus forment le système de recommandation de *Netflix*, comme le décrivent leurs concepteurs. Toutefois, ils soulignent que bien d'autres encore entrent en compte, comme des algorithmes qui caractérisent la « sélection évidente ». Ceux-ci tentent de déterminer ce que l'utilisateur veut vraiment et prennent en compte la popularité des films, leurs métadonnées (les réalisateurs, acteurs, année de réalisation, prix qu'a remporté le film) et génère des images adaptées de chaque film qui peuvent être différentes selon ces croisements d'informations, afin d'améliorer le lancement d'un visionnage [fig.2].

Tous ces algorithmes reposent sur l'apprentissage machine et l'apprentissage statistique.

La complexité de ces systèmes est due à cette multitude de compositions possibles. *Todd Yellin*, vice-président de *Netflix*, dit avoir déjà identifié 2000 communautés de goût en procédant à une étude complète des profils des abonnés. En travaillant avec des « *data scientists* », l'entreprise a pu analyser ce que les internautes consommaient comme séries, et surtout, la fréquence avec laquelle ils passaient d'une série à l'autre.

E. LE FUTUR DES RECOMMANDATIONS

Des chercheurs estiment qu'il leur faut améliorer les recommandations pour ceux qui découvrent *Netflix* durant le premier mois gratuit car le système ne les connaît pas encore. La recommandation destinée aux enfants, en prenant en compte d'une manière prédictive leurs changements de goûts avec l'âge est aussi un défi de l'équipe des développeurs chez *Netflix*. A chaque utilisation, le logiciel algorithmique de la plateforme évolue et devient plus intelligent et plus pertinent. Outre le genre de la série que regarde un abonné, on peut également supposer que *Netflix* tente de caractériser chacune d'entre elles par le lieu dans lequel se déroule l'action, les personnalités des héros, les actions des différents personnages, l'intrigue, la manière dont sont montés les plans, l'atmosphère de la série, la musique, etc. Tous ces critères permettraient d'affiner le chemin virtuel que la plateforme construit pour chacun de ses abonnés, de manière à savoir à qui proposer de nouvelles séries et de prévoir au mieux le comportement des internautes.

III. SOURCES

Netflix, comment fonctionne l'algorithme des recommandations ?

Guillaume Belfiore, Journaliste High Tech, Futura Tech
<http://heig.ch/dadipe>

Apprentissage automatique
Wikipedia, 2012

<http://heig.ch/dizu>

Tout est recommandation

Hubert Guillaud, Internetactu.net, 2017

<http://heig.ch/butatzu>

Les algorithmes de recommandation

Matthieu, Podcastscience.fm, 2012

<http://heig.ch/tazo>

Netflix dévoile comment son algorithme vous rend accros

Romane Mugnier, Usbek & Rica, 2017

<http://heig.ch/besosa>

Netflix Recommendations : Beyond the 5 starts

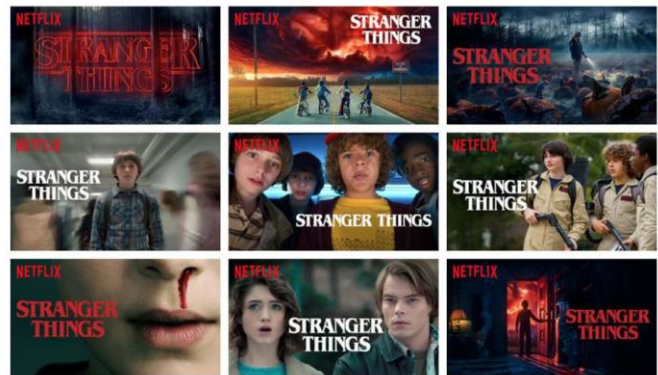
Xavier Amatriain et Justin Basilico (Personalization Science and Engineering), 2012

<http://heig.ch/bibore>

How Recommendation Engines Help You in Your Personalization Efforts, Trouvus, 2015

<http://heig.ch/moza>

IV. ILLUSTRATIONS



[Fig. 1] Illustration du mécanisme de recommandation personnalisée prenant en compte des événements saisonniers. NETFLIX/MEDIUM



[Fig. 2] Illustration du mécanisme de recommandation personnalisée prenant en compte le comportement de l'utilisateur. NETFLIX/MEDIUM