

Intelligence artificielle et publicité

Par Romain NICCOLI
et Franck LE OUAY
Cofondateurs de Criteo

La publicité digitale a permis l'essor de nouveaux géants technologiques, tels Google et Facebook, et l'intelligence artificielle y a joué un rôle clé. À une moindre échelle, c'est aussi la clé du succès de Criteo.

Avant d'examiner plus en détail l'apport de l'IA dans la publicité, commençons par préciser l'objectif de la publicité. Pour simplifier, on peut dire qu'il s'agit de générer des ventes pour des produits ou des services. On distingue généralement les campagnes publicitaires de type *branding*, qui visent à développer l'image d'une marque et dont on attend des revenus supplémentaires, mais sur le long terme, et les campagnes dites de *direct response* ou « DR », qui visent à générer des revenus à court terme, directement liés à la campagne elle-même. Dans la catégorie *branding*, on trouvera, par exemple, un spot publicitaire à la télévision pour Pepsi pendant le Superbowl aux États-Unis : le succès de la campagne se mesurera sur la perception de la marque Pepsi par les consommateurs au travers de sondages, avec l'idée qu'une image positive conduira à des ventes en magasins plus tard. Au contraire, un catalogue papier envoyé par la poste, avec des coupons de réduction, sera à classer dans la catégorie DR : on mesurera le succès de cette action marketing au chiffre d'affaires supplémentaire généré pendant la période de validité des coupons.

Dans les deux cas se pose la question de l'efficacité de la publicité : est-ce que l'impact est à la hauteur de l'argent investi ? Pour répondre à la question, encore faut-il pouvoir mesurer l'impact et ensuite l'optimiser. John Wanamaker (1838-1922), un pionnier du *marketing*, disait déjà à son époque : "Half the money I spend on advertising is wasted; the trouble is I don't know which half". La publicité digitale (sur Internet) a changé la donne, au moins dans le domaine du DR, car elle a rendu l'impact mesurable à un niveau très fin. Il est en effet possible de connaître les revenus correspondant à des publicités vues par chaque internaute, en temps réel. Cela permet d'optimiser les campagnes et de mesurer les changements immédiatement. Cette boucle de rétroaction, couplée avec des quantités de données gigantesques, est un terrain de jeux parfait pour l'intelligence artificielle, qui peut effectuer l'optimisation beaucoup mieux qu'un humain tant il y a de paramètres et de combinaisons à prendre en compte. En effet, puisque la technologie le permet, il est intéressant de personnaliser l'expérience publicitaire pour chacun des millions d'internautes, afin d'en maximiser l'impact. Rien ne sert de montrer une publicité pour des couches à quelqu'un qui n'a pas d'enfant en bas âge. Il faut comprendre les centres d'intérêt de chacun. C'est une vraie révolution par rapport à ce qu'il était possible de faire sur les canaux classiques, comme la presse, la radio ou la télévision.

Le type d'IA qui est à l'œuvre dans la publicité est le *machine learning* ou apprentissage par la machine. Il s'agit d'algorithmes qui digèrent des quantités colossales d'indices sur les interactions passées entre les internautes et les publicités qu'ils ont vues pour apprendre ce qui sera le plus efficace à l'avenir. Il ne s'agit pas vraiment d'intelligence au sens classique du terme, mais plutôt de traitements statistiques très sophistiqués et redoutablement efficaces.

Des algorithmes et des hommes

Creusons un peu cette notion d'intelligence artificielle et en particulier la répartition des tâches entre l'algorithme et l'intelligence humaine. Dans le domaine de la publicité, l'intelligence artificielle est un serviteur dévoué, aux capacités de calcul exceptionnelles un peu comparables à celles du personnage autiste joué par Dustin Hoffman dans le film *Rain Man*, mais elle ne prend pas d'initiatives. Elle fait ce que l'on lui dit de faire et il reste à l'intelligence humaine la responsabilité de définir l'objectif. Il est donc crucial de choisir le bon objectif et le sujet est plus épineux qu'il n'y paraît. Au premier abord, on peut se dire qu'il faut maximiser les ventes à budget publicitaire donné. Plusieurs questions se posent alors.

La question de l'attribution

La première difficulté consiste à relier une vente sur un site de e-commerce à la campagne publicitaire qui en est à l'origine. Et là les choses se corsent. Imaginons une situation où un site de e-commerce mène de front différentes campagnes publicitaires : spots à la télévision, à la radio, affichages de bannières publicitaires sur Internet, etc. Une personne peut se retrouver exposée à plusieurs de ces campagnes. Si, en fin de compte, elle achète un produit, comment savoir laquelle de ces campagnes est à l'origine de cette vente ? À ce jour, personne n'a complètement résolu cette question, dite de l'attribution. L'industrie publicitaire utilise des règles d'attribution arbitraires, où l'on prend en compte l'interaction entre l'internaute et la publicité, et la chronologie des événements. Pour des bannières publicitaires affichées sur Internet, une règle assez répandue consiste à attribuer une vente à la dernière bannière publicitaire sur laquelle l'internaute a cliqué : c'est la règle dite du *post click*.

La question de la fenêtre d'attribution

Toujours concernant la problématique de l'attribution, faut-il prendre en compte les ventes réalisées dans les 24 heures qui suivent un clic sur la publicité, dans les 7 jours, 30 jours ? À nouveau, il n'y a pas de règle absolue et le chiffre choisi va changer énormément la valeur perçue de la publicité. On aimerait mesurer sur une période longue pour prendre en compte la valeur sur le long terme d'un nouveau client, mais l'on sent bien que plus une vente est éloignée du moment où la publicité a été consommée, plus il est douteux de penser que cette vente a été générée par la publicité en question. De plus, il faudrait attendre trop longtemps pour mesurer l'impact de la publicité et l'optimiser. On se limite donc généralement à des durées de l'ordre de quelques jours.

La question de l'incrémentalité

Dans la même veine, si un internaute achète un produit après avoir été exposé à de la publicité, on peut se demander s'il n'aurait pas acheté ce produit tout de même en l'absence de publicité, ou bien s'il l'aurait acheté un peu plus tard. En d'autres termes, le chiffre d'affaires généré par la publicité tel qu'il est mesuré avec la règle d'attribution *post-click* est-il vraiment incrémental ? On peut se rassurer en faisant des tests (dits A/B tests) où l'on sépare la population en deux groupes (A et B), l'un étant exposé à la publicité et l'autre non, pour comparer le chiffre d'affaires total sur chacun des deux groupes. La différence est alors l'impact réel de la publicité. Cela permet de s'assurer que l'impact réel est proche de la valeur mesurée, mais c'est un test statistique sur une grande population, cela ne permet pas d'optimiser une campagne publicitaire à l'échelle de chaque internaute.

Ces quelques exemples permettent de toucher du doigt la complexité de la mesure du succès d'une campagne publicitaire et donc de la définition de ses objectifs. Et nous n'avons même pas parlé de stratégies *marketing* à plus haut niveau : faut-il se concentrer sur ses clients actuels pour les fidéliser et leur faire acheter plus ou sur la conquête de nouveaux clients, par exemple ? Aujourd'hui, l'IA ne

répond pas à ces questions. Elle se contente d'exécuter au mieux la publicité, une fois les objectifs et mesures de succès définis par des humains. C'est déjà beaucoup !

Exemples de fonctionnement de l'IA

Le Search Engine Marketing

Google a été un pionnier de la publicité sur Internet. En octobre 2001, le moteur de recherche introduit un système de publicité qui allait faire sa fortune : AdWords. L'idée est toute simple. Il est désormais possible à un publicitaire d'insérer un petit texte promotionnel dans la colonne à droite des résultats de recherche. Il suffit d'indiquer la liste de « mots-clés » à cibler et Google se chargera d'afficher la publicité pour les internautes concernés. C'est le début de la publicité ciblée sur Internet. La pertinence des textes publicitaires est sans précédent. En effet, tapez « appareil photo » et vous ne verrez certainement que des publicités pour des appareils photos. Mais l'innovation de Google ne s'arrête pas là. Il introduit un *business model* révolutionnaire. Au lieu de se rémunérer sur l'affichage de chaque publicité, Google propose à l'annonceur de ne payer que si l'internaute clique sur la publicité. C'est le « Coût par clic » (CPC). L'espérance de gain d'une publicité est simplement $CTR * CPC$, CTR étant le *Click-Through Rate* ou « Taux de clic », en français. L'estimation précise de ce CTR devient cruciale pour optimiser le placement des publicités. S'il y a 20 annonceurs en compétition pour le mot-clé « couches », mieux vaut afficher celui qui aura le CTR le plus élevé. Cela optimisera le revenu de Google et maximisera également l'intérêt de l'internaute.

Un problème difficile

Comment estimer le CTR sur des publicités correspondant à des millions de mots-clés différents ? Il est fort à parier que le CTR va dépendre également de nombreux paramètres, tels que le mot-clé en question, la position de la publicité sur la page, le texte de la publicité, le « lien » entre la publicité et le mot-clé tapé par l'internaute, la notoriété du site annonceur, la période de l'année, etc. Il y a des milliards de combinaisons distinctes. Mais, comme si le problème n'était pas assez difficile, les taux de clics observés sont relativement faibles (de l'ordre du pourcent), ce qui fait que les événements « positifs », source de rémunération, sont relativement rares. Si nous voulions « mesurer » expérimentalement les CTR, il faudrait des dizaines de milliers d'impressions pour chaque combinaison de paramètres.

L'IA à la rescousse

Il se trouve que ce type de problème entre parfaitement dans le cadre de l'apprentissage supervisé. Cette branche du *machine learning* étudie les problèmes d'apprentissage guidés par des exemples. Chaque affichage constitue un exemple :

- positif (= 1), si l'utilisateur a cliqué ;
- négatif (= 0), si l'utilisateur ne clique pas.

L'algorithme peut être entraîné sur les millions d'exemples avec l'objectif de réduire l'erreur entre la prédiction et la réalité. Il suffit d'enregistrer tous les affichages de publicité et tous les clics pour « nourrir » l'algorithme en continu de nouveaux exemples. L'algorithme peut ainsi naturellement s'adapter à de nouvelles publicités, de nouveaux mots ou même des modes passagères.

L'IA et la publicité graphique

Les techniques de ciblage initialement développées sur les moteurs de recherche ont ensuite trouvé un développement considérable dans la publicité dite « graphique », c'est-à-dire tous les encarts publicitaires que l'on trouve sur Internet. Aujourd'hui, une très grande partie des publicités que vous voyez tous les jours sont achetées aux enchères. Le site qui propose un emplacement va l'offrir au

plus offrant en temps réel. Des dizaines de sociétés spécialisées, comme Criteo, ont alors quelques dizaines de millisecondes pour miser et espérer gagner cette impression. Le gagnant gagne le droit d'afficher sa publicité en direction de l'internaute. L'IA devient un formidable avantage concurrentiel si l'on est capable d'estimer l'espérance de gain de chaque impression de manière très fine et... en temps réel ! Non seulement il est primordial d'estimer des mises très rapidement, mais il faut également mettre à jour ces estimations en temps réel. Imaginez par exemple que vous montrez une publicité à un internaute. Celui-ci ne clique pas. Deux secondes plus tard, vous avez l'opportunité de lui montrer la même publicité. Sachant qu'il n'a pas cliqué la première fois, l'espérance de gain du second affichage en est sûrement diminuée. Les algorithmes d'IA sont très performants et donnent d'excellents résultats dans ces conditions pourtant très difficiles.

L'IA et la personnalisation de la publicité

Au-delà de prédire combien miser pour une opportunité d'affichage, il peut être pertinent également de générer directement des messages publicitaires en fonction des intérêts de l'internaute. Cette technique a notamment été popularisée par Criteo. Avec l'IA, tout devient paramétrable et configurable en temps réel et selon le contexte. Par exemple, on va pouvoir optimiser le nombre de produits montrés dans une bannière publicitaire, les produits eux-mêmes, les photos représentant les produits, les textes, les couleurs, le type d'animation, la présence ou non de prix ou de coupons de réduction, etc. L'IA permet toutes les folies et toutes les optimisations tant qu'il y a des données pour la nourrir. C'est pour cela que, au-delà des algorithmes d'IA, la bataille entre les géants du Web se situe sur l'acquisition de données propriétaires. Si Facebook connaît mes amis et la ville où j'habite, il aura peut-être un avantage compétitif sur le ciblage publicitaire par rapport à un concurrent qui n'aurait pas cette information.

L'IA, un outil universel

Aujourd'hui, l'IA est partout et ses champs d'application ne font que grandir de jour en jour. Tous les fils d'actualité sur Twitter ou Facebook sont mis à jour à partir des interactions de millions d'internautes. Les résultats de Google sont eux-mêmes améliorés en continu avec le *feedback* implicite des utilisateurs. Des campagnes d'emailing sont personnalisées et pilotées par des algorithmes similaires.

L'IA est présente pour reconnaître qu'un même internaute est en fait présent sur deux ordinateurs différents ou pour détecter une fraude.

L'IA permet aussi d'adopter des stratégies d'enchères très complexes, par exemple en n'estimant pas uniquement l'espérance de gain d'un affichage, mais la manière dont vont miser les concurrents potentiels.

Que peut-on attendre de l'IA au service de la publicité dans le futur ?

L'IA continue de faire des progrès spectaculaires, notamment avec le *deep learning*. Ces techniques sont désormais appliquées au service de la publicité pour affiner toujours plus la qualité des messages présentés et les montants à miser sur les places de marché.

Au-delà des progrès des algorithmes existants, il reste une grande marge de progression pour prétendre remplacer complètement les humains. L'IA pourrait apprendre par exemple à mesurer et optimiser le vrai impact de la publicité, en termes de ventes incrémentales sur toute la durée de vie d'un client, plutôt que de se contenter de mesures imparfaites mais plus faciles à réaliser. Ensuite, elle pourrait définir elle-même les stratégies *marketing* de haut niveau, y compris dans l'univers du *branding*. Elle aura alors complètement changé la nature du travail du directeur marketing, mais ce jour paraît aujourd'hui encore assez lointain.