



LA CHRONIQUE DE
GILLES DOWEK

DES ALGORITHMES INSTINCTIFS

Les algorithmes d'apprentissage automatique suscitent une révolution en informatique. Mais il est difficile d'« expliquer » pourquoi un algorithme de ce type donne tel ou tel résultat.



Jean-Michel Basquiat, artiste new-yorkais iconique des années 1980, a laissé une œuvre unique, que l'on reconnaît en un coup d'œil sans y penser. C'est cet « instinct » que reproduisent les algorithmes d'apprentissage profond.

L'adjectif « instinctif » fait partie de ces mots du langage courant que tout le monde comprend mais que personne ne sait définir avec rigueur et qui sont, de ce fait, souvent écartés du vocabulaire scientifique. Pourtant, le développement des algorithmes d'apprentissage automatique (*machine learning*) pourrait nous inciter à donner une seconde vie à cet adjectif.

Pour comprendre pourquoi, faisons une expérience: choisissons deux peintres, par exemple Ambrogio Lorenzetti et Jean-Michel Basquiat, et montrons à un groupe d'enfants deux tableaux de chacun de ces artistes, en leur indiquant à chaque fois qui en est l'auteur. Soumettons-leur ensuite un cinquième tableau et demandons-leur de l'attribuer à un peintre ou à l'autre. Il est probable qu'ils trouvent aussitôt la réponse. Demandons-leur ensuite comment ils ont fait ou, ce qui revient au même, de décrire la méthode qui permet d'attribuer une œuvre à un peintre

ou à un autre. Il est probable qu'ils restent muets.

Cette situation est à comparer avec cette autre: demandons aux mêmes enfants d'additionner les nombres 98 et 99. Non seulement ils trouveront le résultat, mais ils sauront expliquer comment ils ont fait: ils ont ajouté le 8 et le 9, posé le 7, retenu le 1, etc.

Nous ne savons pas programmer un ordinateur pour qu'il distingue un chat d'un chien

Ces deux situations sont très différentes. Dans le cas de l'opération arithmétique, la résolution du problème est consciente – les enfants savent comment ils ont fait. Dans celui de l'attribution d'un tableau, elle est inconsciente – ils ne savent pas comment ils ont fait. De plus,

pour effectuer l'addition, les enfants ont manipulé des symboles: 8, 9, 7, 1... Il semble que ce ne soit pas le cas lorsque nous attribuons un tableau à un peintre – et s'il y a des symboles, nous ne savons pas lesquels.

Pendant longtemps, cette différence a été à l'origine d'un paradoxe: nous savions programmer un ordinateur pour qu'il additionne deux nombres d'un million de chiffres – ce qui, pour un être humain, est très difficile –, mais nous ignorions comment programmer un ordinateur pour qu'il distingue un chien d'un chat – ce qui, pour un être humain, est très facile.

Les algorithmes d'apprentissage automatique ont résolu ce paradoxe. Une méthode parmi d'autres pour distinguer les œuvres de Lorenzetti et de Basquiat consiste à définir une « distance » entre les images, c'est-à-dire d'associer, à chaque paire d'images, un nombre d'autant plus grand que les images sont différentes. Pour attribuer tel tableau à un peintre, il suffit de calculer sa distance à toutes les œuvres de Lorenzetti et à toutes les œuvres de Basquiat d'une base de données: si le tableau est statistiquement plus proche des œuvres de Basquiat, il a de grandes chances d'avoir été réalisé par l'artiste new-yorkais.

Ces algorithmes ne manipulent pas de symboles, sinon les chiffres qui expriment les nombres mesurant les distances, et, comme les enfants de l'expérience ci-avant, nous avons du mal à expliquer pourquoi ces programmes attribuent une œuvre à un peintre plutôt qu'à un autre. Ils apportent un argument statistique, mais ils ne sont pas en mesure de justifier leur résultat.

Nous avons sans doute besoin d'un adjectif pour qualifier ces algorithmes non symboliques et non conscients, afin de les distinguer des autres, tel l'algorithme de l'addition. Or, il existe en français un adjectif qui qualifie les impulsions souvent irraisonnées qui déterminent nos actes: l'adjectif « instinctif ». ■

GILLES DOWEK est chercheur à l'Inria et membre du conseil scientifique de la Société informatique de France.