

Les algorithmes et leurs biais dans le viseur des mouvements anti-racistes

Un afro-américain arrêté à tort à cause d'un logiciel de reconnaissance faciale : cette affaire a relancé le débat sur les biais de l'intelligence artificielle, en pleine mobilisation contre le racisme et les violences policières.

Temps de lecture : minute

7 juillet 2020

L'affaire remonte à début janvier : Robert Williams est arrêté à Detroit et passe 30 heures en détention parce qu'un logiciel a jugé identique la photo de son permis de conduire et l'image d'un voleur de montres capturée par des caméras de surveillance. À tort. Pour l'Union américaine des libertés civiles (ACLU), qui a porté plainte en son nom le 24 juin, *"bien qu'il soit le premier cas connu, il n'est probablement pas la première personne qui a été arrêtée et interrogée à tort sur la base d'une fausse reconnaissance faciale"* .

Pour Joy Buolamwini, fondatrice du groupe d'activistes Algorithmic Justice League, cette affaire est révélatrice *"de la façon dont le racisme systémique peut être encodé et reflété dans les intelligences artificielles (IA)"* . Sous la pression d'associations comme la puissante ACLU, Microsoft, Amazon et IBM avaient d'ailleurs annoncé début juin restreindre l'utilisation de leurs outils d'analyse de visages par les forces de l'ordre.

L'IA est fondée sur l'apprentissage automatisé à partir de données insérées par le concepteur, que la machine analyse. Si ces données sont

biaisées, le résultat est faussé. Une étude du Massachusetts Institute of Technology parue en février 2018 a ainsi révélé une forte disparité selon les groupes de population, avec des taux d'erreur de moins de 1% pour les hommes blancs, et allant jusqu'à 35% pour les femmes noires, dans les principaux logiciels de reconnaissance faciale passés au crible.

Thermomètre et pistolet

Dans un tweet devenu viral, Nicolas Kayser-Bril, de l'ONG Algorithm Watch, montre que confronté à des images d'individus tenant un thermomètre frontal, le programme d'analyse d'images "Google Vision" reconnaissait des "jumelles" dans une main à peau blanche mais identifiait "un pistolet" dans la main à peau noire. Selon lui, ce biais était *"probablement dû au fait que les images utilisées dans la base de données qui comportaient des personnes noires étaient plus souvent associées à la violence, que celles avec des personnes blanches"* .

Google a reconnu auprès d'Algorithm Watch un résultat "inacceptable". Or des logiciels de ce type sont légion, et commercialisés auprès d'entreprises et d'administrations du monde entier, pas seulement par de grands noms de la tech. *"Cela rend très difficile l'identification des conditions dans lesquelles l'ensemble de données a été collecté, des caractéristiques de ces images, de la façon dont l'algorithme a été formé"*, estime Seda Gürses, chercheuse à l'université de technologie de Delft aux Pays-Bas.

Cette multiplicité d'acteurs permet de réduire les coûts, mais cette complexité brouille le traçage et l'attribution de responsabilités, selon la chercheuse. *"Un policier raciste peut être formé ou remplacé, tandis que dans le cas d'un algorithme"*, les décisions dans les entreprises sont déterminées par cet algorithme, qui obéit à des critères avant tout économiques, selon Seda Gürses.

Un constat qui s'applique aussi à des programmes prétendant prédire des comportements criminels. En témoigne une récente polémique autour d'un logiciel se vantant de *"prédire avec 80% de précision si une personne est criminelle en se basant uniquement sur une photo de son visage"*. Plus de 2 000 personnes, dont de nombreux scientifiques, ont signé une pétition demandant à l'éditeur Springer Nature de ne pas publier un article consacré à cette technologie mise au point par plusieurs professeurs de l'université de Harriburg en Pennsylvanie, et défendue dans cet article. Mais l'éditeur, interrogé par l'AFP, a assuré qu'il n'avait *"jamais été accepté pour publication"*. Le problème, ce n'est pas tant l'algorithme que celui des présupposés des chercheurs, commente Nicolas Kayser-Bril, qui récuse l'approche *"purement technologique"* .

"Même avec des jeux de données d'excellente qualité, on n'arrive à rien si on ne prend pas en compte toutes les problématiques sociales qu'il y a derrière. Pour cela il faut travailler avec des sociologues, a minima".

"Vous ne pouvez pas changer l'histoire du racisme et du sexisme, estime Mutale Nkonde, chercheuse en intelligence artificielle pour les universités de Stanford et Harvard. Mais vous pouvez faire en sorte que l'algorithme ne devienne pas le décideur final".