

Google veut rendre le machine learning moins discriminant

Le sexisme, le racisme et d'autres formes de discrimination peuvent s'immiscer dans les algorithmes du machine learning qui rendent nos systèmes « intelligents », comme l'a souligné Kate Crawford de Microsoft Research dans les colonnes du [New York Times](#).

L'an dernier, par exemple, des [chercheurs de l'université Carnegie Mellon](#) ont constaté que les femmes étaient moins susceptibles que les hommes de voir s'afficher des annonces d'emploi au salaire très élevé sur Google.

Dès sa conception, l'algorithme peut donc apprendre des biais de jugements humains (voir [le cas du bot Tay](#)). La firme de Mountain View veut y remédier et rendre l'intelligence artificielle (IA) moins discriminante, avec l'aide de ses chercheurs.

« Égalité des opportunités »

Dans un document technique (« *l'égalité des opportunités dans l'apprentissage supervisé* ») des chercheurs membres de l'équipe Google Brain proposent un cadre pour tester le caractère discriminatoire des algorithmes existants, à partir d'un ensemble d'attributs sensibles. Et « ajuster » les prédicteurs (également nommés variables indépendantes ou variables d'entrée) pour réduire ce risque. Il s'agit de trouver l'équilibre entre la précision et le caractère non-discriminatoire du résultat.

« *Quand l'appartenance au groupe coïncide avec un attribut sensible, comme la race, le genre, le handicap ou la religion, cette situation peut conduire à des résultats injustes ou préjudiciables* », explique dans un [billet de blog](#) Moritz Hardt, co-auteur de la recherche avec Eric Price et Nathan Srebro.

L'approche de Google, celle de « *l'égalité des opportunités* » appliquée à l'apprentissage automatique, résume-t-il, est basée sur l'idée que les individus qui peuvent bénéficier d'un résultat souhaité devraient « *avoir une même chance d'être classés correctement* » pour ce résultat.

Lire aussi :

[Avec Assistant, Google espère supplanter Siri](#)

[Google met de l'intelligence artificielle dans son traducteur](#)

[Une IA pour rendre Google Analytics plus intelligent](#)