

# Machine Learning : Microsoft veut chasser les biais pour réduire les discriminations

«Ce que nous découvrons {dans notre document}, c'est que ce n'est pas parce qu'une machine agit de manière agnostique qu'elle sera impartiale ... L'apprentissage automatique c'est: vous regardez le monde, puis vous apprenez du monde. La machine va donc également apprendre les biais qui existent dans le monde qu'elle observe » explique une équipe de recherche de Boston University qui [a travaillé](#) sur les biais des algorithmes de machine learning, avec le soutien de Microsoft.

Un constat que les spécialistes admettent [depuis longtemps](#) et que l'éditeur de Redmond a décidé de prendre en charge en développant un outil capable de détecter les biais dans les algorithmes d'intelligence artificielle. Objectif : aider les entreprises à [utiliser le machine learning](#) sans courir le risque de discriminer certaines personnes.

## **Des risques de discrimination à l'emploi**

Pour Rich Caruana, qui travaille sur le projet pour Microsoft, il s'agira d'un «tableau de bord» que les ingénieurs pourront appliquer aux modèles d'IA pour corriger les biais natifs.

« Des choses comme la transparence, l'intelligibilité et l'explication sont assez récentes pour que peu d'entre nous aient suffisamment d'expérience pour savoir tout ce que nous devrions rechercher et toutes les façons dont ce biais pourrait se cacher dans nos modèles », explique-t-il au [MIT Technology Review](#) .

Les conséquences peuvent être pénalisantes dans les domaines de la santé, de la justice et de l'emploi notamment, pointent les chercheurs de Boston University. Ainsi, on constate des connexions sémantiques entre « programmeur » et « homme », plus rarement avec le mot « femme ». Un biais qui pourrait provoquer une ségrégation à l'emploi.

« La chose la plus importante que les entreprises peuvent faire maintenant est d'éduquer leur main-d'œuvre afin qu'elles soient conscientes des myriades de biais qui peuvent apparaître et se manifester et créer des outils pour rendre les modèles plus faciles à comprendre » explique-t-il à [Venture Beat](#)

D'autres acteurs de l'IT travaillent sur cette problématique. En début du mois, Facebook a annoncé tester une solution appelée Fairness Flow sur son service d'offres d'emploi pour traquer les biais discriminants. Idem chez IBM qui propulse sa solution IA, Watson, comme [une priorité stratégique](#).

*Crédit Photo : @Merrill College of Journalism Press Releases via Visual Hunt / CC BY-NC*