Machine Learning au service de l'épidémiologie : ciblage des restaurants à inspecter par les services sanitaires aux États-Unis

14 décembre 2018

En novembre 2018 ont été diffusés les résultats de travaux menés, aux États-Unis, par Google, l'université Harvard et les départements de santé et d'innovation de Las Vegas et Chicago, visant à améliorer le ciblage des restaurants à contrôler par les services sanitaires.

L'équipe de recherche a tout d'abord mis en place un algorithme de détection des requêtes, lancées sur le navigateur web Google, concernant des problèmes de santé consécutifs à la consommation d'aliments dans des conditions sanitaires médiocres. L'algorithme permet de distinguer les personnes effectivement malades de celles faisant des recherches dans un autre cadre : par exemple, les médecins et les étudiants peuvent se renseigner sur certains symptômes sans être eux-mêmes atteints. Il vérifie également si les personnes, dont les requêtes ont été identifiées comme pertinentes, sont bien allées au restaurant les jours précédents, ce grâce au service de localisation de Google (sous réserve qu'il n'ait pas été désactivé). Les restaurants sont ensuite classés comme « à risque » lorsqu'ils sont à l'origine d'une proportion importante de requêtes. L'application ainsi construite (FINDER) a été déployée dans les services sanitaires de Chicago (de novembre 2016 à mars 2017) et de Las Vegas (de mai à août 2016).

52,3 % des restaurants, identifiés par FINDER et contrôlés par les services sanitaires, se sont avérés non conformes lors des inspections, contre 24,7 % en temps normal. Comparativement, 39,4 % des restaurants contrôlés à Chicago suite au dépôt d'une plainte se sont révélés non conformes (résultats de Las Vegas non disponibles), soit une efficacité de détection moindre par rapport à FINDER. Ceci s'explique notamment par le fait qu'une personne peut avoir des difficultés à identifier quel restaurant l'a rendue malade, parmi ceux visités.

Ainsi, pour les auteurs, FINDER est un outil prometteur d'aide à la mise en œuvre de politiques sanitaires dans le secteur de la restauration, en améliorant le ciblage des inspections. Cette application présente l'avantage de ne pas nécessiter d'action de la part des consommateurs (par ex. porter plainte). Cependant, elle suppose un accès à leurs données personnelles de localisation et de recherches Google. Concernant le coût de mise en place de l'outil, les services testeurs rapportent qu'un certain effort a été nécessaire au début, mais qu'une fois pris en main, sa maintenance nécessite peu de ressources et de temps, tout en leur fournissant des indications précieuses pour mener leurs enquêtes terrain.

Nombre moyen d'infractions à la réglementation par catégorie (« critique » ou « majeure ») rencontrées dans les restaurants identifiés par FINDER ou non (*Baseline*)

	FINDER* n = 132	BASELINE* n = 5848	p-value
Critical violations	0.40	0.21	0.001
Major violations	0.74	0.56	0.04

*Adjusted mean violation count, accounting for city and risk level, calculated using linear regressions

Source : *npj Digital Medicine*

Aurore Payen, Centre d'études et prospective

Source : [npj Digital Medicine](#)