



SÉMINAIRE DE L'AXE | Médicament et santé des populations



Méthodes d'apprentissage automatique causal pour la détection d'effets de traitement hétérogènes *Causal Machine Learning Methods for Heterogeneous Treatment Effect Detection*

Philippe Boileau

Professeur adjoint, Département d'épidémiologie, de biostatistique et de santé au travail et Département de médecine, Université McGill

Mardi 24 mars 2026 à 11h30 au pavillon Jean-Coutu, S1-127

L'Axe médicament et santé des populations de la Faculté de pharmacie a le plaisir de vous inviter à un séminaire présenté par **Philippe Boileau**, professeur adjoint, Département d'épidémiologie, de biostatistique et de santé au travail et Département de médecine de l'Université McGill.

Biographie

Philippe Boileau, Ph. D., est professeur adjoint en biostatistique à l'Université McGill. Il est également chercheur junior à l'Institut de recherche du Centre universitaire de santé McGill. M. Boileau s'intéresse de façon générale au développement de méthodes statistiques dites "faiblement fondées sur les hypothèses" (assumption-lean) et à leur application à des problèmes quantitatifs en sciences de la santé et de la vie. Ces approches allient l'inférence causale et l'apprentissage machine afin de fournir une inférence statistique fiable sans reposer sur des hypothèses de convenance. Il est aussi engagé dans le développement de logiciels statistiques libres et, plus largement, dans la promotion de pratiques de recherche reproductible. Ses travaux les plus récents portent sur le développement de méthodes d'apprentissage machine causal pour la découverte d'effets de traitement hétérogènes dans les données d'essais cliniques.

Philippe Boileau, PhD is an Assistant Professor of Biostatistics at McGill University. He is also a Junior Scientist at the Research Institute of the McGill University Health Centre. Dr. Boileau is broadly interested in the development of assumption-lean statistical methods and their application to quantitative problems in the health and life sciences. Assumption-lean procedures combine causal inference and machine learning to provide dependable statistical inference without relying on convenience assumptions. He is also committed to the development of open-source statistical software and, more broadly, to the adoption of reproducible research practices. His most recent work has focused on developing causal machine learning methods for heterogeneous treatment effect discovery in clinical trial data.

Résumé

L'effet moyen conditionnel du traitement (conditional average treatment effect [CATE]) est fréquemment estimé afin de réfuter l'hypothèse d'un effet de traitement homogène. Selon cette hypothèse, toutes les unités constituant la population étudiée bénéficient de manière identique d'un traitement donné. Cependant, la mise en évidence d'effets de traitement hétérogènes par inférence sur le CATE exige que les variables modifiant réellement l'effet du traitement soient recueillies de façon fiable au départ. Les techniques basées sur le CATE échoueront nécessairement à détecter des violations lorsque des modificateurs d'effet sont omis des données, par exemple en raison de contraintes de ressources. Une erreur de mesure importante a un impact similaire. Pour pallier ces limites, nous démontrons que l'hypothèse d'un effet de traitement homogène peut être évaluée par inférence sur les contrastes des variances des résultats potentiels (potential outcomes). Nous dérivons aussi des estimateurs d'apprentissage automatique causal pour ces contrastes. Ces procédures d'inférence sont ensuite utilisées pour détecter des effets de traitement hétérogènes lors de la réanalyse d'un essai contrôlé randomisé portant sur la gestion ciblée de la température chez des patients en arrêt cardiaque.

The conditional average treatment effect (CATE) is frequently estimated to refute the homogeneous treatment effect assumption. Under this assumption, all units making up the population under study experience identical benefit from a given treatment. Uncovering heterogeneous treatment effects through inference about the CATE, however, requires that covariates truly modifying the treatment effect be reliably collected at baseline. CATE-based techniques will necessarily fail to detect violations when effect modifiers are omitted from the data due to, for example, resource constraints. Severe measurement error has a similar impact. To address these limitations, we prove that the homogeneous treatment effect assumption can be gauged through inference about contrasts of the potential outcomes' variances. We also derive causal machine learning estimators of these contrasts. These inference procedures are then used to detect heterogeneous treatment effects in the re-analysis of a randomized controlled trial investigating targeted temperature management in cardiac arrest patients.