

Projet de recherche

Doctorante: Céline Béji

Etablissement: MILES, Pôle Sciences des données, LAMSADE, Université Paris Dauphine-PSL, CNRS.

Directeur de thèse: Ph. Jamal Atif, Florian Yger

Sujet: Revisiter le problème d'inférence causale à la lumière des nouvelles avancées en apprentissage automatique.

Date: Mars 2017 à aujourd'hui

La causalité est centrale dans de nombreuses applications, comme en médecine, en économie et en marketing [1]. Dans le développement de la médecine personnalisée par exemple, pouvoir cibler les catégories de personnes qui répondent positivement à un traitement médical, c'est-à-dire pour lesquelles le rétablissement est "causé" par le traitement administré, aurait des conséquences non-négligeables en santé publique.

Dans la continuité des travaux de Rubin qui aborde la causalité comme un problème à données manquantes [2], nous reformulons l'inférence causale comme une estimation de densité. Prenons l'exemple où nous voudrions évaluer l'efficacité d'un traitement médical sur un groupe de patients malades. Sois $X \in \mathfrak{M}_{n,p}(\mathbb{R})$ une matrice représentant n patients chacun caractérisé par p attributs, comme leur âge, leur poids, leurs antécédents médicaux et leurs indicateurs de suivi. Notons $Y_i(1)$ la variable illustrant la réaction du patient au traitement si le traitement est administré à l'individu i et $Y_i(0)$ s'il ne l'est pas. À partir des réactions observées dans chacun des deux cas, quatre populations causales peuvent être identifiées. La plus intuitive est celle pour laquelle le traitement est bénéfique. Les patients se rétablissent s'ils reçoivent le traitement et restent malades dans le cas contraire ($Y_i(1) = 1, Y_i(0) = 0$). Certains patients peuvent se rétablir par eux-mêmes, avec ou sans traitement, ($Y_i(1) = 1, Y_i(0) = 1$), ou au contraire rester malades quelle que soit la situation ($Y_i(1) = 0, Y_i(0) = 0$). Ces individus constituent les deuxièmes et troisièmes groupes identifiés. Enfin, en raison des effets indésirables, des patients qui aurait été convalescents sans traitement peuvent se voir indisposés sous l'effet du traitement. Le but de ma thèse est d'identifier ces populations causales, afin d'une part de déterminer les patients pour lesquels le traitement a eu un impact positif et d'autre part de mettre en place une politique d'affectation du traitement pour de nouveaux patients. Contrairement aux tests traditionnels qui évaluent l'efficacité d'un traitement médical en moyenne, il serait possible de cibler les patients positifs au traitement même s'ils ne représentent qu'une minorité de la population. Le problème fondamental réside dans le fait qu'il est expérimentalement impossible d'observer pour un même patient sa réaction avec ou sans traitement.

Notre première contribution a été de modéliser l'inférence causale comme un problème d'estimation de densité de probabilité d'un mélange de populations distinctes; chacune étant définie par la réaction potentielle avec et sans traitement. Nous introduisons comme information partielle dans notre modèle des contraintes basées sur la structure causale des données. Par exemple, pour un patient s'étant vu administrer le traitement et y ayant répondu positivement ($Y_i(1) = 1$), deux scénarios sont

envisageables : soit le traitement présente l'effet attendu, soit l'évolution du patient aurait été positive dans tous les cas. Ceci exclu les deux autres choix de population auxquels l'individu pourrait appartenir. Leur probabilité de distribution est alors fixée à zéro. Nous avons pu démontrer que l'estimation de ces distributions permettait d'évaluer l'effet individuel du traitement ainsi que sa moyenne sur tous les individus. Enfin, nous avons étendu l'interprétation des populations causales au cas multi-traitements, permettant de prendre en compte l'effet combiné des traitements.

Une deuxième contribution a consisté à implémenter une adaptation de l'algorithme EM [3] qui estime les paramètres des distributions des populations tout en tenant compte des contraintes de causalité. Nous avons démontré l'existence d'une borne inférieure pour la log-vraisemblance ainsi que la convergence de l'algorithme vers un maximum global. Cette deuxième assertion est produite par l'introduction des contraintes de causalité qui induisent une identification unique des distributions estimées à chacune des populations causales. Son efficacité a été expérimentalement prouvée pour sur une mixture de Gaussienne. De plus, nous avons démontré que sur des jeux de données contenant des variables catégorielles, une mixture de Gaussiennes et de Multinomiales indépendantes améliorée les résultats. Un algorithme variationnelle dérivé de celui-ci a également été proposé pour pallier la non-dérivabilité directe de la vraisemblance marginale.

Enfin, nous proposons une approche non-paramétrique. Nous utilisons un auto-encodeur [4] enrichit d'un a priori causal, matérialisé par un masque appliqué sur la couche intermédiaire de notre réseau. Les caractéristiques des individus sont reconstruites après avoir été réduites sur l'espace latent caractérisant les populations causales. Lors d'expérimentations, nous discutons le nombre de neurones à adresser sur la couche latente et démontrons l'efficacité de notre méthode par rapport aux modèles de références sur données synthétiques et réelles.

Notre approche a l'avantage d'être applicable en multi-traitement. Elle permet d'obtenir un intervalle de confiance sur la prédiction des résultats de chaque traitement ainsi que de déterminer la combinaison optimale des traitements qui maximise l'effet global. Dans nos travaux futurs, il est envisagé d'étudier la corrélation entre les différents traitements, qui reste, à ce jour, inexplorée.

References

- [1] Judea Pearl, Madelyn Glymour, and Nicholas P Jewell. *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.
- [2] D.B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974.
- [3] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977.
- [4] Yoshua Bengio et al. Learning deep architectures for ai. *Foundations and trends® in Machine Learning*, 2(1):1–127, 2009.