CentraleSupélec

Inférence Statistique des Relations de Causalité

Vers des algorithmes d'apprentissage éthique?

Encadrant : Frédéric Pennerath

Mohammed FELLAJI, Ahmed BEN AISSA

September, 2020

Contents

1 Introduction

2 Bibliographic study

one of the main papers : [1] link

2.1 let's start with an example

2.2 Causation vs Association

[1] [2] [3] One of commun phrases in statistics is

2.3 Association - Intervention - Counterfactuals

[1]

2.4 Paper : Judea Pearl

see [3]
One of the best paper so far. It explains clearly the difference between the causal concept and the associational concept (for example : correlation). The first concept ...

2.5 Notations

see [4] and [2]

2.6 definition of counterfactual

2.7 Simpson's paradoxe

- give definition :

- give example : Cholesterol

2.8 Average Causal Effect

see [2]

3 Definition of the causality

Perhaps the most important message of the discussion and methods presented in this paper would be a widespread awareness that (1) all studies concerning causal relations must begin with causal assumptions of some sort and (2) that a friendly and formal language is currently available for articulating such assumptions.[3]

[5]

4 Python Library : DoWhy

library developed by Microsoft [6], blog article available link

5 Causal discovery

A traditional way to discover causal relations is to use interventions or randomized experiments, which is, however, in many cases of interest too expensive, too time- consuming, unethical, or even impossible. Therefore, inferring the underlying causal structure from purely observational data, or from combinations of observational and experimental data, has drawn much attention in various disciplines. With the rapid accumulation of huge volumes of data, it is necessary to develop automatic causal search algorithms that scale well.[7]

References

[1] J. Pearl, "The seven tools of causal inference, with reflections on machine learning," *Communications of the ACM*, vol. 62, no. 3, pp. 54–60, 2019.

[2] M. A. Hernán and J. M. Robins, "Causal inference: what if," *Boca Raton: Chapman & Hill/CRC*, vol. 2020, 2020.

[3] J. Pearl, "The mathematics of causal relations," *Causality and Psychopathology: Finding the Determinants of Disorders and their Cures (P. Shrout, K. Keyes and K. Ornstein, eds.). Oxford University Press, Corvallis, OR*, pp. 47–65, 2010.

[4] L. Yao, Z. Chu, S. Li, Y. Li, J. Gao, and A. Zhang, "A survey on causal inference," *arXiv preprint arXiv:2002.02770*, 2020.

[5] D. B. Rubin, "Causal inference using potential outcomes: Design, modeling, decisions," *Journal of the American Statistical Association*, vol. 100, no. 469, pp. 322–331, 2005.

[6] "DoWhy: A Python package for causal inference." https://github.com/microsoft/dowhy, 2019.

[7] C. Glymour, K. Zhang, and P. Spirtes, "Review of causal discovery methods based on graphical models," *Frontiers in Genetics*, vol. 10, p. 524, 2019.

[8] A. Marx and J. Vreeken, "Testing conditional independence on discrete data using stochastic complexity," *arXiv preprint arXiv:1903.04829*, 2019.

[9] B. Schölkopf, "Causality for machine learning," *arXiv preprint arXiv:1911.10500*, 2019.

[10] J. Pearl *et al.*, "Causal inference in statistics: An overview," *Statistics surveys*, vol. 3, pp. 96–146, 2009.

[11] P. Judea, "Causality: models, reasoning, and inference," *Cambridge University Press. ISBN 0*, vol. 521, no. 77362, p. 8, 2000.

[12] J. Pearl, M. Glymour, and N. P. Jewell, *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.

[13] J. Pearl, "The science and ethics of causal modeling," *Handbook of ethics in quantitative methodology*, pp. 383–416, 2011.

[14] H. Matute, F. Blanco, I. Yarritu, M. Díaz-Lago, M. A. Vadillo, and I. Barberia, "Illusions of causality: how they bias our everyday thinking and how they could be reduced," *Frontiers in Psychology*, vol. 6, p. 888, 2015.