

Causal Machine Learning

Krikamol Muandet

CISPA Helmholtz Center for Information Security

Society consists of diverse individuals, demographic groups, and institutions. Developing and deploying algorithmic models across these varied environments involves navigating numerous trade-offs. To create reliable machine learning algorithms capable of effective real-world interaction, addressing this heterogeneity is essential. In particular, the ability to identify and leverage causal relationships is a critical component of building reliable AI systems.

In this tutorial, I will introduce the fundamentals of causal inference within the context of machine learning. I will begin by highlighting the mutual importance of causality in machine learning. Next, I will discuss how instrumental variables (IVs) can be used to mitigate the impact of unobserved confounders, enhancing the credibility of algorithmic decision-making and the reliability of models built from observational and heterogeneous data. Specifically, I will demonstrate how we can use machine learning tools, such as kernel methods and deep learning, to address potentially ill-posed problems in non-linear IV regression and proxy variable applications. Finally, I will emphasize that understanding how data is generated and how models can influence it is crucial for reliable human-machine interactions, especially when complete information about the data may be inaccessible.