



Practical Causal Reasoning as a Means for Ethical ML

Isabel Valera

(Saarland University in Saarbrücken)

25/06/2025, 4:15 pm

Department of Statistics, Ludwigstr. 33, Seminar Room 144,
and online via Zoom [\[Link\]](#)
(Meeting-ID: 631 1190 7291; Password: StatsCol)

Abstract

In this talk I will give an overview of the role of causality in ethical machine learning, and in particular, in fair and explainable ML. In particular, I will first detail how to use causal reasoning to study fairness and interpretability problems in algorithmic decision making, stressing the main limitations that we encounter when aiming to address these problems in practice. Then, I will introduce the audience to causal generative models, a novel class of deep generative models that do not only accurately fit observational data but can also provide accurate estimates to interventional and counterfactual queries. I will focus this part of the talk on our recent paper DeCaFlow, a deconfounding causal generative model (CGM). DeCaFlow can provably identify all causal queries with a valid adjustment set or sufficiently informative proxy variables. Remarkably, for the first time to our knowledge, we show that a confounded counterfactual query is identifiable, and thus solvable by DeCaFlow, as long as its interventional counterpart is as well.

About the Speaker:

Isabel Valera is Full Professor of Machine Learning at the Department of Computer Science at Saarland University, and Adjunct Faculty at the MPI for Software Systems in Saarbrücken. She is the recipient of an ERC Starting Grant on “Society-Aware ML”, and a Fellow of the European Laboratory for Learning and Intelligent Systems (ELLIS). Her research focuses on the development of trustworthy machine learning methods that can be used in the real world. Her research can be broadly categorized into three main themes: fair, interpretable, and robust machine learning. Her research interests cover a wide range of ML approaches, including deep learning, probabilistic modeling, causal inference, time series analysis, and many more.

References:

Almodóvar, A., Javaloy, A., Parras, J., Zazo, S. and Valera, I., 2025. DeCaFlow: A Deconfounding Causal Generative Model. *arXiv preprint arXiv:2503.15114*.
