



DATA SCIENCE INITIATIVE
BROWN UNIVERSITY



FLAVIO DU PIN CALMON

*Assistant Professor of Electrical Engineering
John A. Paulson School of Engineering
Harvard University*

On Representations and Fairness: A Few Information-Theoretic Tools for Machine Learning

Information theory can shed light on the algorithm-independent limits of learning from data and serve as a design driver for new machine learning algorithms. In this talk, we discuss a set of information-theoretic tools that can be used to (i) help understand fairness and discrimination in machine learning and (ii) characterize data representations learned by complex learning models. On the fairness side, we explore how local perturbations of distributions can help both identify proxy features for discrimination as well as repair models for bias. On the representation learning side, we explore a theoretical tool called principal inertia components (PICs), which enjoy a long history in the statistics and information theory literature. We use the PICs to scale-up a multivariate statistical tool called correspondence analysis (CA) using neural networks, enabling data dependencies to be visualized and interpreted at a large scale. We illustrate these techniques in both synthetic and real-world datasets and discuss future research directions.

Flavio P. Calmon is an Assistant Professor of Electrical Engineering at Harvard's John A. Paulson School of Engineering and Applied Sciences. Before joining Harvard, he was the inaugural data science for social good post-doctoral fellow at IBM Research in Yorktown Heights, New York. He received his Ph.D. in Electrical Engineering and Computer Science at MIT. His main research interests are information theory, inference, and statistics, with applications to fairness, privacy, machine learning, and communications engineering. Prof. Calmon has received the NSF CAREER award, the Google Research Faculty Award, the IBM Open Collaborative Research Award, and Harvard's Lemann Brazil Research Fund Award.

Wednesday, April 24

4:00 PM – 5:00 PM

164 Angell Street, 3rd Floor Open Space

Coffee and pastry to be served. Hosted by Sarah M. Brown.
Sponsored by the Data Science Initiative.