

Yuekai Sun, Assistant Professor of Statistics

Title: Does enforcing fairness mitigate algorithmic biases due to distributional shift?

Abstract: Many instances of algorithmic bias are caused by distributional shifts. A particularly prominent class of examples is algorithmic biases caused by under-representation of samples from minority groups in the training data. We study whether enforcing algorithmic fairness during training mitigates such biases in the target domain. On one hand, we show that there are scenarios in which enforcing fairness does not improve model performance (in the target domain). In fact, it may even harm performance. On the other hand, we derive sufficient conditions under which enforcing group and individual fairness successfully mitigates algorithmic biases due to distributional shifts.

Bio: Dr. Yuekai Sun is an assistant professor in the statistics department at the University of Michigan. His research is guided by the statistical and computational challenges in machine learning. Some topics of recent interest are:

- algorithmic fairness,
- federated learning,
- learning in non-IID settings.

More broadly, Yuekai is interested in the mathematical foundations of data science. He obtained his PhD in computational mathematics from Stanford University, where he worked with Michael Saunders and Jonathan Taylor.

Photo:

