

Optimizing Society? Ensuring Fairness in Automated Decision-Making

Sorelle Friedler

Department of Computer Science, Haverford College
Haverford, Pennsylvania, USA
sfriedle@haverford.edu

Abstract

Algorithms are increasingly used to make high-stakes decisions about people; who goes to jail, what neighborhoods police deploy to, and who should be hired for a job. But if we want these decisions to be fair, this means we must take societal notions of fairness and express them using the language of math. What is a fair decision and how can it be guaranteed?

In this talk, we'll discuss recent work from the new and growing field of Fairness, Accountability, and Transparency. We will examine technical definitions of fairness and non-discrimination that have been proposed and their societal counterparts. We'll also discuss methods for ensuring that algorithms are making decisions as desired, from methods to audit black-box algorithms to white-box interpretability techniques. This important field necessitates societally informed and mathematically rigorous work; we'll discuss open problems in this light.

2012 ACM Subject Classification Security and privacy → Social aspects of security and privacy

Keywords and phrases algorithmic fairness

Digital Object Identifier 10.4230/LIPIcs.SWAT.2018.2

Category Invited Talk



© Sorelle Friedler;

licensed under Creative Commons License CC-BY

16th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2018).

Editor: David Eppstein; Article No. 2; pp. 2:1–2:1

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany