Egalitarian justice and machine learning

ML and Economic Inequality
April 19-20th 2021

Reuben Binns
Dept. of Computer Science, University of Oxford, United Kingdom
reuben.binns@cs.ox.ac.uk

Outline

- Egalitarian justice in philosophy
- How has egalitarianism featured in 'fair' machine learning?
- Does egalitarianism 'fit' ML?
- Dimensions of egalitarian justice:
  - Object of justice
  - Site of justice
  - Responsibility for justice
  - Ideal vs non-ideal theory
  - The namespace of justice

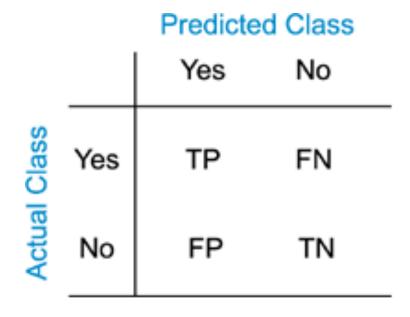# How has egalitarianism featured in 'fair' machine learning?

- Early work (e.g. Dwork et al 2012) referenced egalitarian political philosophy including Rawls and Roemer
- But the measures of fairness in ML typically focus on avoiding direct or indirect discrimination as defined in law (typically, US law)

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012, January). Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference* (pp. 214-226).

Egalitarian justice in fair machine learning

- Outcome parity
- Equalised error rates
- Calibration

Classification error parity

Calibration

Lessons from political philosophy

- If we care about egalitarian justice in ML, political philosophy may help
- Internally diverse but offer a guide
- Equality of what?
- Choice and luck
- Deontic vs telic justice

Binns, R. (2018, January). Fairness in machine learning: Lessons from political philosophy. In *Conference on Fairness, Accountability and Transparency* (pp. 149-159). PMLR.

Egalitarian justice in fair machine learning
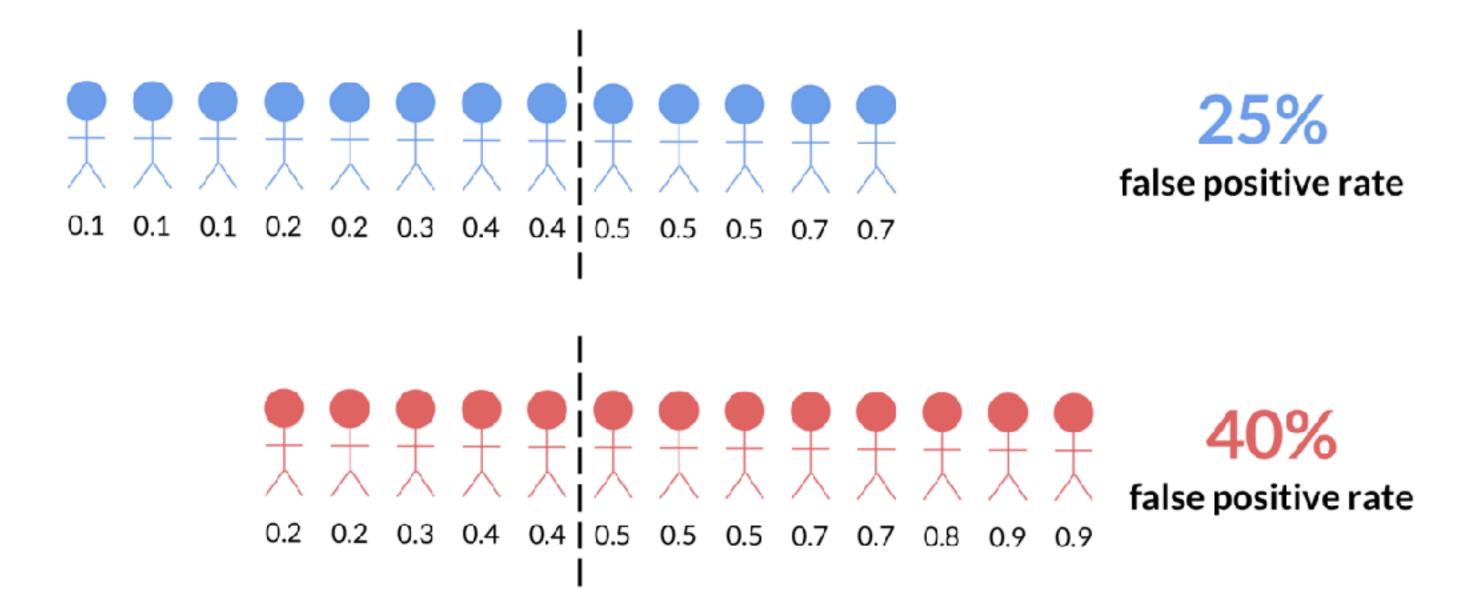
- Outcome parity
- Equalised error rates
- Calibration

Egalitarian justice in fair machine learning

- Heidari et al 2019 - luck egalitarian equality of opportunity

- Counterfactual - a kind of limited luck egalitarianism

- Kasy & Abebe 2021 - causal impact of algorithm on inequality

- Friedler et al 2016/2021 - depends on worldview assumptions

- Barabas 2018 - interventions to prevent compounding bad luck

Heidari, Hoda, et al. "A moral framework for understanding fair ML through economic models of equality of opportunity." *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 2019.

Kasy, M., & Abebe, R. (2021, March). Fairness, equality, and power in algorithmic decision-making. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 576-586).

Friedler, Sorelle A., Carlos Scheidegger, and Suresh Venkatasubramanian. "The (Im) possibility of fairness: different value systems require different mechanisms for fair decision making." *Communications of the ACM* 64.4 (2021): 136-143.

Barabas, Chelsea, et al. "Interventions over predictions: Reframing the ethical debate for actuarial risk assessment." *Conference on Fairness, Accountability and Transparency*. PMLR, 2018.

Embedding egalitarian justice in an ML model?

- Can and should we apply egalitarian justice to ML models? How?
- Should we be taking theories of overarching political principles and building them into constraints on classifiers?
- We need to consider additional dimensions:
  - The object of justice
  - The site/subject of justice
  - Responsibility for justice
  - Ideal vs non-ideal theory
  - The gamespace of justice

The object of justice

- What's the correct currency of justice?
- Concerns for egalitarian ML:
  - Is accuracy a primary good to be 'distributed'?
  - What implications do particular currencies have for real world ML?
  - Relational vs distributive, power

The site/subject of justice

- Where is justice supposed to happen? What kinds of relationships give rise to its demands?
- Rawlsian 'basic structure'; interpersonal; cosmopolitan global structures
- Micro and macro inequalities
- Fallacies of composition

Responsibility for justice

- Who is responsible for correcting injustice, e.g. redistributing wealth
- State and social institutions? Interpersonal?
- Responsibility and structural injustice

Ideal vs non-ideal theory

- What is the point of thinking about ideal theory when the world is not ideal?
- Mills' non-ideal theory
- Are we trying to characterise injustice, or figure out how to address it?

Gamespace of justice

- E.O. Wright's 'gamespace' (Wark 2019)
  - Situational
  - institutional
  - systemic
- What level are we applying egalitarian justice to?

Level of System at which Conflict Is Focused	Game Metaphor	Political Form of Conflict	Stakes in the Conflict	Form of Class Analysis
System level	What game to play	Revolutionary versus counter-revolutionary	Capitalism versus socialism	Marxist
Institutional level	Rules of the game	Reformist versus reactionary	Varieties of capitalism	Weberian
Situational level	Moves in the game	Interest group politics	Immediate economic interests	Durkheimian

Table 6.1. The Game Metaphor for Mapping Politics and Class Analysis

Thanks! Questions?

Reuben Binns
Dept. of Computer Science, University of Oxford, United Kingdom
reuben.binns@cs.ox.ac.uk