Justice, Fairness, Equality and Machine Learning

Samuel Freeman

Three Kinds of Justice:
1. Formal Justice

- 'Treat similar cases similarly'

- Most basic precept of the rule of law

- Means the administration of laws should apply equally, in the same way, to those belonging to the classes defined by laws.

- Also connected with Idea of Impartiality.

2. Substantive Justice

- Formal justice is not a guarantee of the substantive justice of laws themselves. Unjust laws can be impartially administered.

- E.g. the laws of a caste society that discriminates against racial or religious classes.

- Substantive justice involves respect for persons as moral and social equals.

- It is realized by guarantee of equal basic rights and liberties: freedom of conscience, expression, association; freedom of the person; freedom of occupation, equality of opportunity, etc.   May involve some idea of distributive justice= a fair distribution of income and wealth.

3. Procedural Justice

- Requires impartial judgment and treating like cases alike, but also making decisions that observe relevant constraints and do not violate individuals' equal rights and liberties.
- Almost all government decision procedures are imperfect (e.g. civil and criminal trials) and can result in false outcomes. Aim is to design decision procedures so that they are more likely to result in the correct decisions. That seems to be one of the primary reasons to incorporate machine learning into government decisions.

Imperfect Procedural Justice

- But considerations of fairness to persons also must be built into these imperfect procedures, sometimes without regard to their effect on outcomes of the procedures. Partly that is a requirement of substantive justice in a democratic society, where people are to be respected as equals.
- Does Machine Learning raise any special problems here—of perpetuating while obscuring unfairness, in the form of discrimination, etc.?

Publicity of laws and their reasons

- Another problem is accessibility, transparency, and the comprehensibility of the outcomes and the reasons for them, that are arrived at by machine learning algorithms.

- One requirement of a democracy is *publicity, of both laws and the reasons behind coercive political decisions*. That we be able to understand the reasons why government makes decisions and exercises its coercive power in the way it does is said to be a condition of individual freedom and autonomy in a liberal-democratic society.

- This is connected with idea that we should be able to *justify to one another in terms of public reasons that all can understand and accept the use of coercive political authority*. Can "rule-making by robot" meet that condition, and if not is that a problem?