

A Left Realist Critique of the Political Value of Adopting Machine Learning Systems in Criminal Justice

Fabio Massimo Zennaro¹

Department of Informatics
University of Oslo

¹fabiomz@ifi.uio.no

Fairness of **adopted** machine learning systems



Fairness of **adopting** machine learning systems

*Fairness of **adopted** machine learning systems.*
How does my system perform in terms of fairness metrics?
What fairness metrics make sense?



*Fairness of **adopting** machine learning systems.*
What is the role of a ML system?
What is the political value of a ML system?

What is the *political value* of a *ML system*?

ML system

- A *supervised system*
- modelling a *functional relationship* between input and output
- by minimizing a given *loss function*

Political value

- Artifacts (and technology) are *not neutral* [4, 3]
- Values it *expresses* (*implicit*)
- Values it makes easy to *express* (*instrumental*)

What is the political value of adopting ML system in the context of criminal justice?

Criminal justice

- Parole assessment, risk assessment, recidivism assessment, police deployment, crime prevention...
- Long historical relationship with statistics [1]
- Very relevant in fair ML [5, 2]

How to uncover political values in the adoption of ML systems in criminal justice?

Critical assessment through the lens of Left Realism (LR)

- Criminological theory from 1980s proposed by *Lea and Young* [4]
- Middle ground between *left idealism* and *law and order* (L&O)
- Often define in opposition to L&O

We assess where ML systems lie on the spectrum of concerns between LR and L&O with respect to six issues.

1. Understanding of Crime

L&O



LR

- Priority in fighting crime
- Reliance on simplified sociological category
- ML focus on correlation and effects
- ML coarse categorical labeling
- Priority in understanding crime
- Complex explanation of behaviour
- (Casual ML)

∴ ML systems may be used to disregard cause-effects links and entrench oversimplification of criminological categories

2. Types of Crime

L&O



LR

- Focus on specific crimes
- Authoritative definition of crime
- Wider view on crimes (“white-collar”)
- Awareness of gap between definition and perception
- ML reliance on standardized data
- ML strict labeling
- (Transfer ML)

∴ ML systems may reinforce historical data and definition given by labeling

3. Data interpretation

L&O



LR

- Direct use of statistics
- Statistics for enforcing policies
- Deep analysis of statistics
- Careful use in decision-making
- ML statistical assumption
- ML as a decision-making tool
- (Statistical ML)

∴ ML systems may foster an instrumental-legalistic processing of data

4. Policing

L&O



LR

- Military policing
- Unilateral enforcement
- Consensus policing
- Community integration
- ML as unilateral data analysis
- (...)

∴ ML systems may better support military policing approaches

5. Accountability

L&O



LR

- Police efficiency
- Police secrecy
- Democratic overview
- Transparency
- ML as hardly interpretable
- ML as black box
- (Interpretable ML)

∴ ML systems may favour opaque policies on the ground of efficiency

6. Analogy with CCTV

1980s

- Adoption of *CCTV*
- New technology
- Promise of efficiency
- Idea of data intelligence
- Idea of remote control

2010s

- Adoption of *ML systems*
- New technology
- Promise of efficiency
- Idea of data intelligence
- Idea of remote control

∴ ML systems may promote a “fire-brigade” mentality in law enforcement

Simplified analogical analysis of ML systems:

- ML systems are more than *functional models*
- Criminal justice theory *richer* than a simple dichotomy

Yet, ML systems have *implicit political biases* that may be overlooked or exploited

- Danger of *naive adoption* (technological enthusiasm)
- Danger of *instrumental adoption* (justification of aims through ML means)

Adopting ML systems is a choice that has a *political value*.

Development of technology ha made the adoption of technology more political [4].

Thanks

Thank you for listening!

References I

- [1] Chelsea Barabas, Karthik Dinakar, Joichi Ito Virza, Jonathan Zittrain, et al. Interventions over predictions: Reframing the ethical debate for actuarial risk assessment. *arXiv preprint arXiv:1712.08238*, 2017.
- [2] Richard Berk, Hoda Heidari, Shahin Jabbari, Michael Kearns, and Aaron Roth. Fairness in criminal justice risk assessments: the state of the art. *arXiv preprint arXiv:1703.09207*, 2017.
- [3] Luciano Floridi. Infraethics—on the conditions of possibility of morality. *Philosophy & Technology*, 30(4):391–394, 2017.
- [4] John Lea, Jock Young, et al. What is to be done about law and order? 1984.
- [5] Rashida Richardson, Jason Schultz, and Kate Crawford. Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review Online*, *Forthcoming*, 2019.