Fair policy learning from observational data

Dennis Frauen & Stefan Feuerriegel

Institute for AI in Management
www.ai.bwl.lmu.de

Agenda

- Motivation
- Problem setting
- Fair policy learning

Motivation

- Automatic decision-making is widespread
- **Examples**: Hiring, credit lending, personalized advertising

However, automated tools can adopt **biases** from historical data[1]:
- Gender-based discrimination in automated hiring
- Race bias in algorithmic risk scoring (COMPAS[2])

Amazon ditched AI recruiting tool that favored men for technical jobs

Specialists had been building computer programs since 2014 to review résumés in an effort to automate the search process

Example: Data-driven hiring process

CV

Gender, race

Hiring policy

Utility

!

How to learn a **new policy**, taking into account **ethics** or **laws**?

1) De-Arteaga M, Feuerriegel S, Saar-Tsechansky M (2022) Algorithmic fairness in business analytics: Directions for research and practice. POM
2) https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis

Related work: 3 building blocks

Algorithmic fairness
How to incorporate sensitive attributes into a classifier?

Off-policy learning/ Causal machine learning
How to evaluate a policy from observational data without introducing bias?

Fairness for resource allocation
How to allocate a limited amount of resources fairly?

Fair policy learning from observational data
We learn fair policies with clever use of tools from causal machine learning!

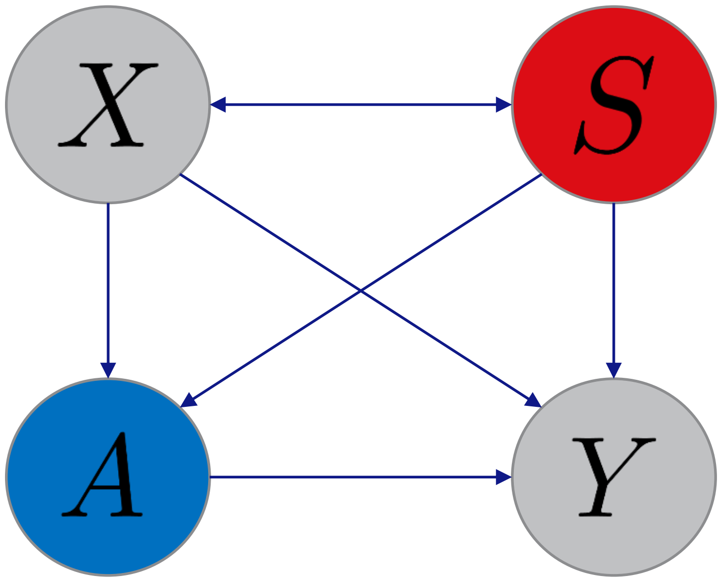
Agenda

- Motivation
- Problem setting
- Fair policy learning

Problem setting

Unsensitive covariates (experience, education)

Sensitive covariates (gender, race)



Treatment (hiring decision)

Outcome (benefit metric)

- Input: observational data $(X_i, S_i, A_i, Y_i)_{i=1}^n$
- Notation: $Y(a)$ denotes the (potential) outcome under treatment intervention $A = a$

Policy learning

- A policy π assigns a an individual with covariates X, S to a probability of receiving treatment $\pi(X, S)$
- Policy value: $V(\pi) = \mathbb{E}[Y^{\pi}] = \mathbb{E}[\pi(X,S)Y(1) + (1 - \pi(X,S))Y(0)]$
- **Goal**: Find policy that maximizes the policy value: $\pi^* \in \arg\max_{\pi \in \Pi} V(\pi)$
- Causal identification: The policy value can be estimated from data, **if all confounders are observed**. $\Rightarrow$ Sensitive confounders **cannot be ignored** in policy value estimation.

Agenda

- Motivation
- Problem setting
- Fair policy learning

Fairness criterea for off-policy learning

1) Action fairness
Policy recommendations should not depend on sensitive covariates

$$\pi(X, S) \perp\!\!\!\perp S$$

2) Value fairness
Policy learning algorithm should take policy value $V_s(\pi)$ **conditioned on the sensitive covariate** into account

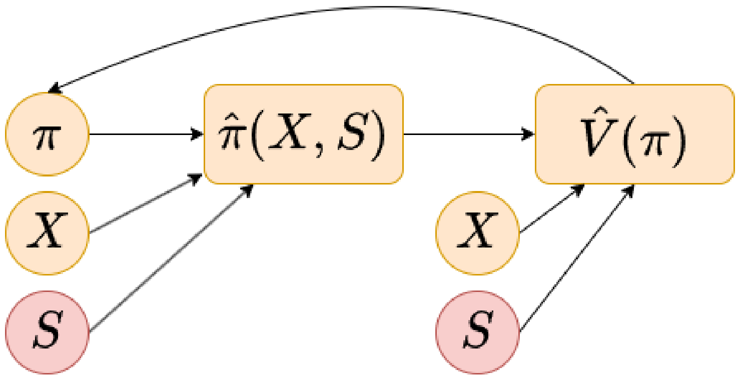
$$V_s(\pi) = E[Y^\pi | S = s]$$

Fairness issues in off-policy learning

Two sources of „unfairness“:

1. Policy depends explicitely on S / covariates correlated with S
2. Policy value is an expectation over X, S and depends on the distribution of S on the observed data

   ⇨ Removing S leads to **unobserved confounding** and **biased estimates**



Why value fairness? Toy example

- S = Gender, X = Age independent, policy only depends on X
- Treatment benefits males, harms females
- 80% of the population is male

  ⇨ Policy will always tend to treat (depending on Age)

  ⇨ Policy value will be larger for males than females

Fairness criterea for off-policy learning

1) Action fairness
Policy recommendations should not depend on sensitive covariates

$$\pi(X, S) \perp\!\!\!\perp S$$

2) Value fairness
Policy learning algorithm should take policy values for each sensitive attribute into account

$$V_s(\pi) = E[Y^\pi | S = s]$$

Max-min fairness

$$\arg\max_{\pi \in \Pi} \min_{s \in S} \hat{V}_s(\pi)$$

Envy-free fairness

$$\arg\max_{\pi \in \Pi} \hat{V}(\pi)$$

$$|V_s(\pi) - V_{s'}(\pi)| \leq \alpha \text{ for all } s, s' \in S$$

Proposed causal machine learning method

Value fair policy objectives

Envy-free

$$\hat{\pi}^{\lambda} \in \arg\max_{\pi\in\Pi} \hat{V}_{\lambda}(\pi)$$
$$\hat{V}_{\lambda}(\pi) = \hat{V}(\pi) - \lambda \max_{s,s'\in\mathcal{S}} |\hat{V}_s(\pi) - \hat{V}_{s'}(\pi)|$$

Max-min

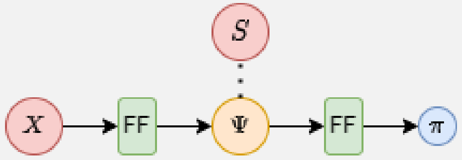
$$\hat{\pi}^{mm} \in \arg\max_{\pi\in\Pi} \min_{s\in\mathcal{S}} \hat{V}_s(\pi)$$

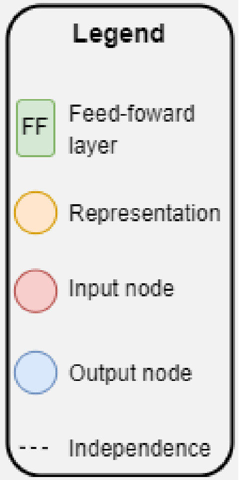
Neural framework for action fair & value fair policies

Neural network parametrization π_θ, loss $\mathcal{L}(\theta)$

Envy-free

$$\mathcal{L}(\theta) = -\hat{V}_{\lambda}(\pi_\theta) + \alpha\mathcal{W}(\Phi(\theta), S)$$





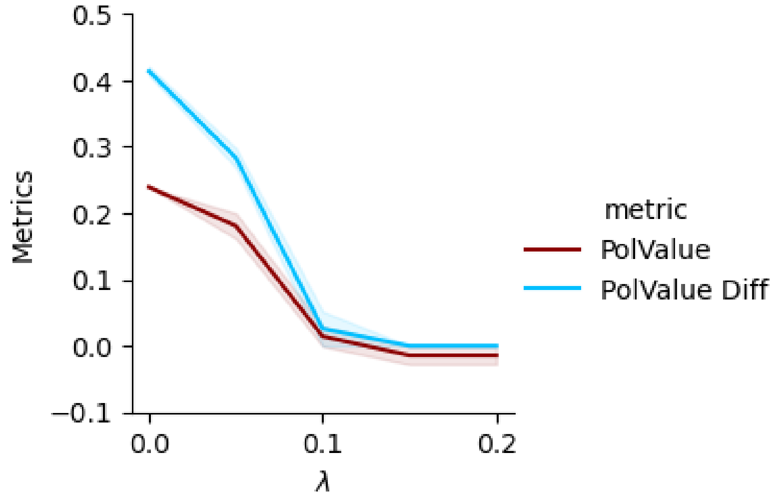
Max-min

$$\mathcal{L}(\theta) = -\min_{s\in\mathcal{S}} \hat{V}_s(\pi_\theta) + \alpha\mathcal{W}(\Phi(\theta), S)$$

Experimental results: simulated data

	Unfair	Action fair + no value fairness	Action fair + max-min fair	Action fair + Envy-free $\lambda = 0.05$
Policy value	0.35 ± 0.04	0.20 ± 0.06	0.19 ± 0.05	0.15 ± 0.03
Policy value (S = 0)	0.17 ± 0.05	0.08 ± 0.04	0.11 ± 0.04	0.07 ± 0.02
Policy value (S = 1)	0.54 ± 0.06	0.45 ± 0.14	0.39 ± 0.08	0.34 ± 0.04

Policy value and difference in policy values between sensitive groups plotted over envy-free parameter λ



Current state of research

Summary	✓ Fairness concepts for policy learning from observational data ✓ Deviation of finite sample estimators for value-fair policies ✓ Generalization bounds ✓ Novel neural framework that learns action fair & value fair policies via representation learning ✓ Experiments using simulated + real-world data
Implications	✓ Novel application of causal machine learning ✓ We tackle fairness issues for policy learning

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

Dennis Frauen
Institute of AI in Management
frauen@lmu.de