Causality-based Fair Machine Learning for Sequential Decision Making

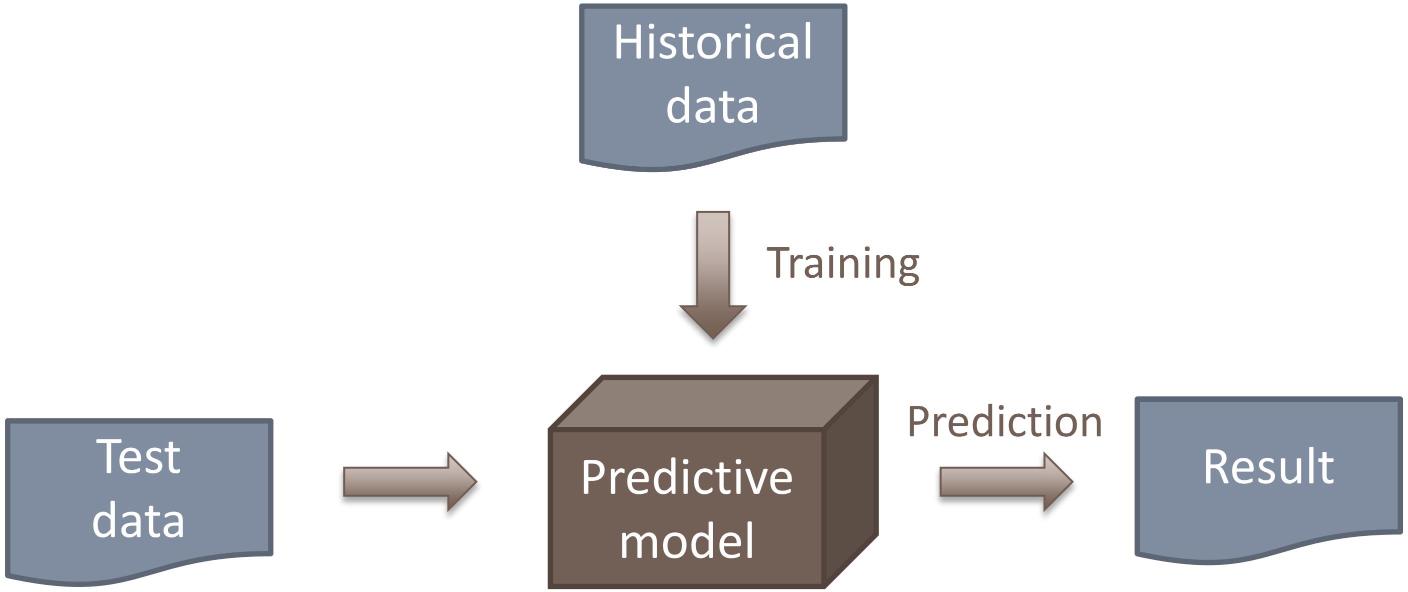
Dr. Lu Zhang

Assistant Professor at University of Arkansas

Fair Machine Learning

- Discrimination refers to unjustified distinctions of individuals based on their membership in a certain group.
- Federal Laws and regulations disallow discrimination on several grounds:
  - Gender, Age, Marital Status, Race, Religion or Belief, Disability or Illness ……
  - These attributes are referred to as the protected attributes.

Fair Machine Learning



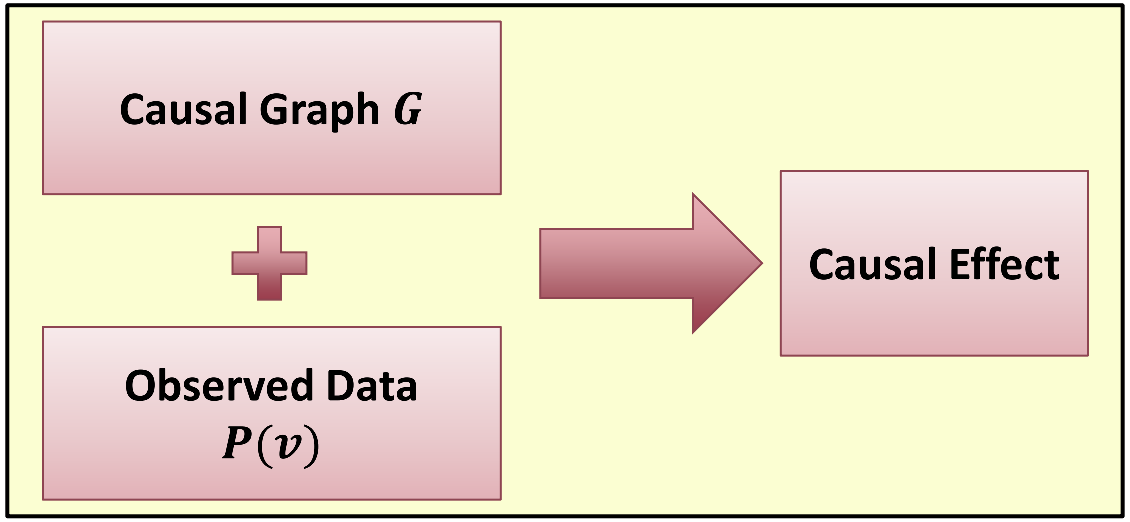
Build discrimination-free predictive decision model

Static vs. Sequential Setting

- Most studies are based on static settings where one-shot decisions are made on tasks such as classification and regression.
- In practical situations, decision-making is more of a sequential nature.
  - Decision models are deployed and make decisions sequentially.
  - Data arrive and are observed sequentially.
- Fair machine learning in the sequential setting is underexplored.
  - Multiple decision models connected in a partial order.
  - Single decision model executed repeatedly and creates feedback loops.
  - Online recommendation where customers arrive in a sequential manner.

Background

- Structural Causal Model (Pearl, 1995): A mathematical framework for describing the causal mechanisms of a system as a set of structural equations.
- Describe how causal relationships and causal effects can be inferred from observed data.



Structural Causal Model

- A causal model is triple $\mathcal{M} =< \boldsymbol{U}, \boldsymbol{V}, \boldsymbol{F} >$, where
  - $\boldsymbol{U}$ is a set of exogenous (hidden) variables whose values are determined by factors outside the model;
  - $\boldsymbol{V} = \{X_1, \cdots, X_i, \cdots\}$ is a set of endogenous (observed) variables whose values are determined by factors within the model;
  - $\boldsymbol{F} = \{f_1, \cdots, f_i, \cdots\}$ is a set of deterministic functions where each f_i is a mapping from $\boldsymbol{U} \times (\boldsymbol{V} \setminus X_i)$ to X_i. Symbolically, f_i can be written as

$$x_i = f_i(\boldsymbol{pa}_i, \boldsymbol{u}_i)$$

where $\boldsymbol{pa}_i$ is a realization of X_i's parents in $\boldsymbol{V}$, i.e., $\boldsymbol{Pa}_i \subseteq \boldsymbol{V}$, and $\boldsymbol{u}_i$ is a realization of X_i's parents in $\boldsymbol{U}$, i.e., $\boldsymbol{U}_i \subseteq \boldsymbol{U}$.

Causal Graph

- Each causal model $\mathcal{M}$ is associated with a direct graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where
  - $\mathcal{V}$ is the set of nodes represent the variables $\boldsymbol{U} \cup \boldsymbol{V}$ in $\mathcal{M}$;
  - $\mathcal{E}$ is the set of edges determined by the structural equations in $\mathcal{M}$: for X_i, there is an edge pointing from each of its parents $\boldsymbol{Pa}_i \cup \boldsymbol{U}_i$ to it.
    - Each direct edge represents the potential direct causal relationship.
    - Absence of direct edge represents zero direct causal relationship.
- Assuming the acyclicity of causality, $\mathcal{G}$ is a directed acyclic graph (DAG).

A Causal Model and Its Graph

Observed Variables $\boldsymbol{V} = \{I, H, W, E\}$

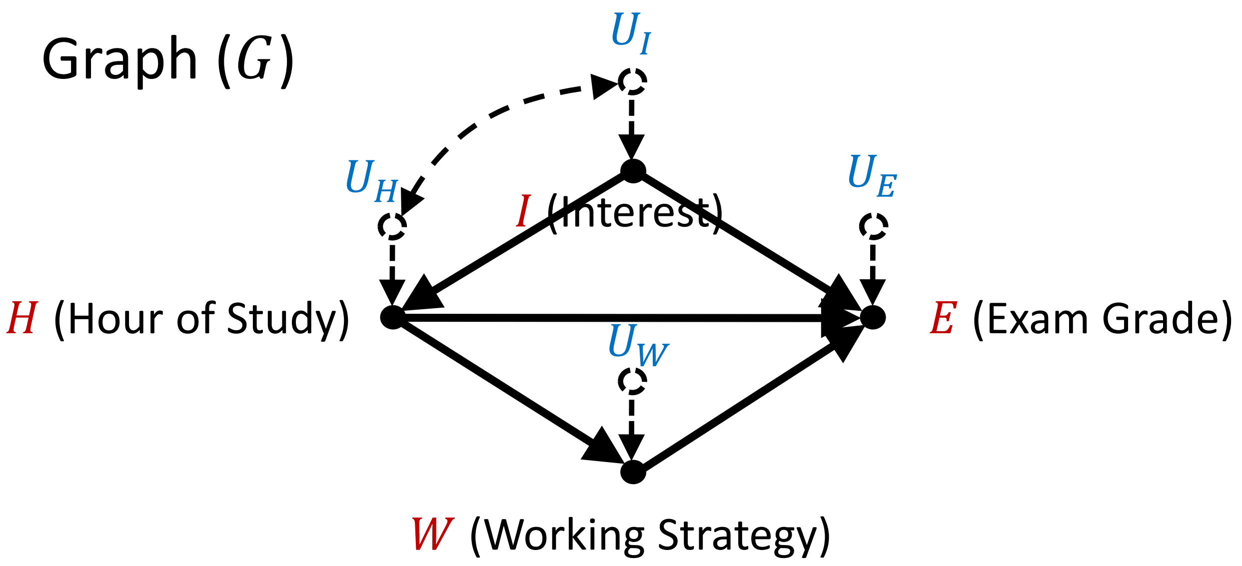
Hidden Variables $\boldsymbol{U} = \{U_I, U_H, U_W, U_E\}$

Model (M)

$$i = f_I(u_I)$$
$$h = f_H(i, u_H)$$
$$w = f_W(h, u_W)$$
$$e = f_E(i, h, w, u_E)$$

Assume I and H are confounded.

Graph (G)



A Markovian Model and Its Graph

Observed Variables $\boldsymbol{V} = \{I, H, W, E\}$

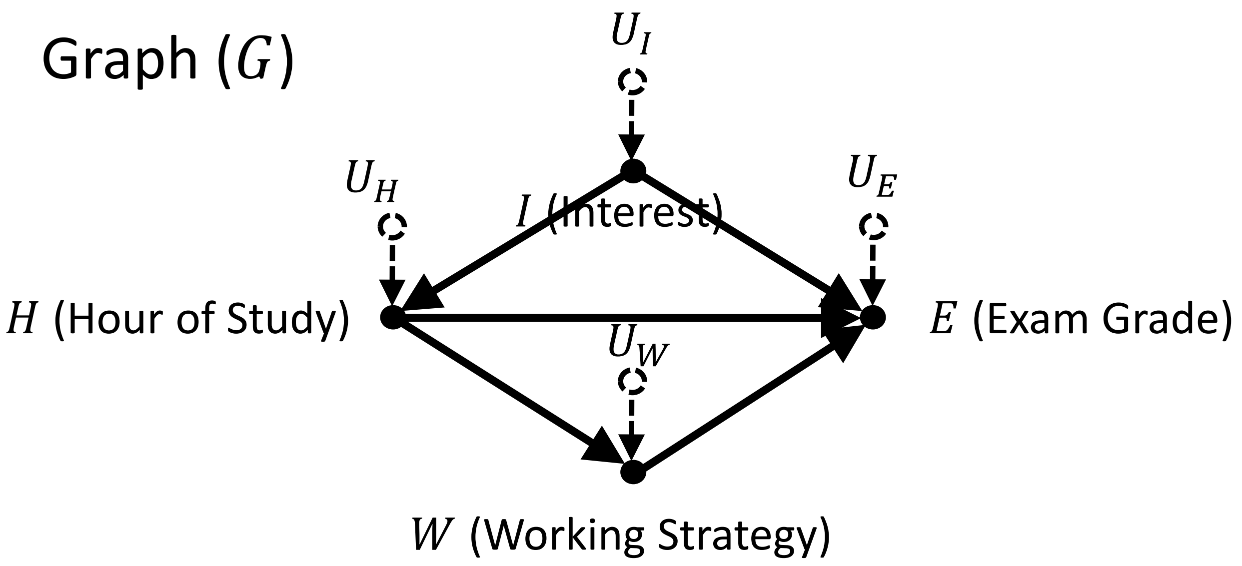
Hidden Variables $\boldsymbol{U} = \{U_I, U_H, U_W, U_E\}$

Model (M)

$$i = f_I(u_I)$$
$$h = f_H(i, u_H)$$
$$w = f_W(h, u_W)$$
$$e = f_E(i, h, w, u_E)$$

Assume no hidden confounder.

Graph (G)



Causal Inference and (Hard) Intervention

- The basic operation of manipulating a causal model.
  - Simulate the manipulation of the physical mechanisms by some physical interventions or hypothetical assumptions.
  - Forces some observed variables $\boldsymbol{X} \in \boldsymbol{V}$ to take certain constants $\boldsymbol{x}$.
- Mathematically formulated as $do(\boldsymbol{X} = \boldsymbol{x})$ or simply $do(\boldsymbol{x})$.
- For an observed variable Y disjoint with $\boldsymbol{X}$, its interventional variant under intervention $do(\boldsymbol{x})$ is denoted by $Y_{\boldsymbol{X} \leftarrow \boldsymbol{x}}$ or $Y_{\boldsymbol{x}}$.
- The effect of intervention on all other observed variables $\boldsymbol{Y} = \boldsymbol{V} \backslash \boldsymbol{X}$ is represented by the post-intervention distribution of $\boldsymbol{Y}$.
  - Denoted by $P(\boldsymbol{Y} = \boldsymbol{y} | do(\boldsymbol{X} = \boldsymbol{x}))$ or simply $P(\boldsymbol{y} | do(\boldsymbol{x}))$;
  - Or equivalently $P(\boldsymbol{Y}_{\boldsymbol{X} \leftarrow \boldsymbol{x}} = \boldsymbol{y})$ or simply $P(\boldsymbol{y}_{\boldsymbol{x}})$.

Soft Intervention

- Force variables to take a functional relationship in responding to some other variables.
- Example: $do\big(Y = h_\theta(\boldsymbol{X})\big)$ is a soft intervention that substitutes structural equation associated with Y by $h_\theta(\boldsymbol{X})$.

- Effect of soft intervention is represented by post-intervention distributions.

- Under certain assumptions (e.g., Markovian), both hard and soft interventional distributions can be computed from observed data.

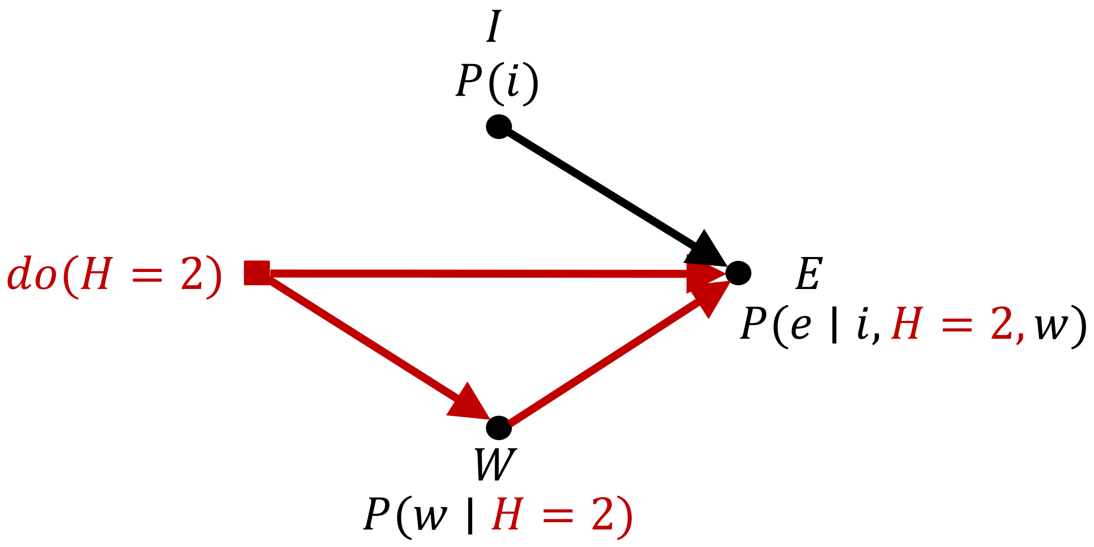
Total (Causal) Effect

- The total effect measures the causal effect transmitted along all causal paths.
- The total effect of X on Y under two interventions $do(x_1), do(x_2)$:

$$TE(x_2, x_1) = P(y|do(x_2)) - P(y|do(x_1))$$

- Example:

$$P(E = \text{‘}A\text{’}|do(H = 2))$$



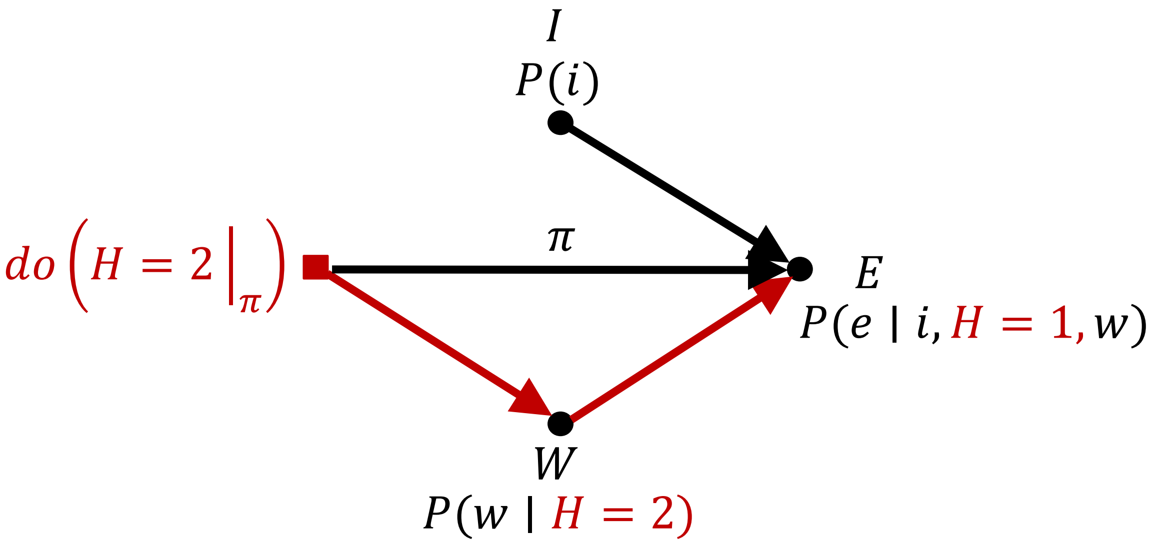
Path-Specific Effect

- Path-specific effect measures the causal effect transmitted along certain causal paths.
- Given a subset of causal paths π, the causal effect of X on Y transmitted along π under two interventions $do(x_1)$ (a.k.a., reference), $do(x_2)$:

$$SE_\pi(x_2, x_1) = P(y|do(x_2|_\pi)) - P(y|do(x_1))$$

- Example:

$$P(E = \mathrm{A}|do(H = 2|_\pi))$$



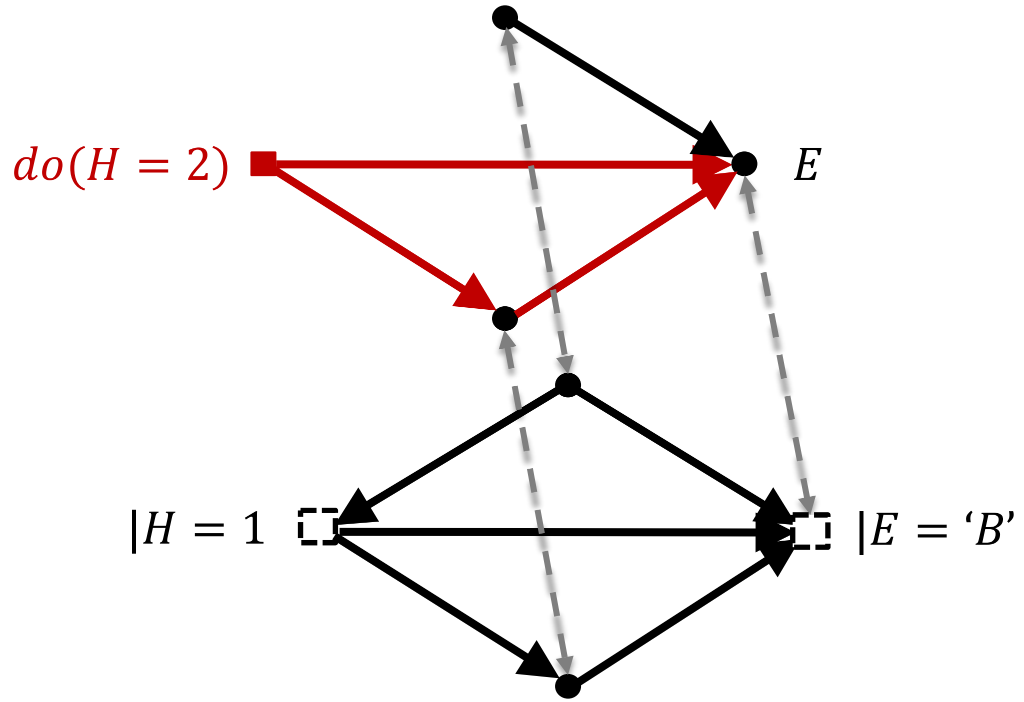
Counterfactual Effect

- Counterfactual effect measures the causal effect while we have certain observations or evidence.
- Counterfactual effect of X on Y under two interventions $do(x_1), do(x_2)$ given that we observe $\boldsymbol{O} = \boldsymbol{o}$:

$$CE(x_1, x_2|\boldsymbol{o}) = P\left(y_{x_1}\middle|\,\boldsymbol{o}\right) - P\left(y_{x_2}\middle|\,\boldsymbol{o}\right)$$

- Example:

$P(E = \text{‘}A\text{’}_{H=2}|H = 1, E = \text{‘}B\text{’})$



Causality-based Fairness Notions

- Protected attribute: $S = \{s^+, s^-\}$
- Profile attributes: X
- Decision: Y

- Demographic parity:
$$TE(s^+, s^-) = 0$$
- Direct/indirect non-discrimination
$$SE_\pi(s^+, s^-) = 0$$
- Counterfactual fairness
$$CE(s^+, s^- | \boldsymbol{x}) = 0$$

FAIR MULTIPLE DECISION MAKING

Y. Hu, Y. Wu, L. Zhang, and X. Wu. Fair Multiple Decision Making through Soft Interventions. NeurIPS, 2020.

Problem Setting

- Consider multiple decision models such that:
  - One decision model may influence one another;
  - Feature distribution may change due to deployment of decision models;

- All decision models may contain discrimination, either be introduced by themselves or transmitted from upstream models.

- Objective: Build fair models for all decision-making tasks.

Challenge

- Building a fair model for each task independently may not work.

- Toy example:

$$X_1 \rightarrow Y_1 \rightarrow X_2 \rightarrow Y_2$$

Toy Example: $X_1 \rightarrow Y_1 \rightarrow X_2 \rightarrow Y_2$

$(X_1, Y_1) \quad (X_2, Y_2)$

Step 1: data collection

Toy Example: $X_1 \to Y_1 \to X_2 \to Y_2$

$(X_1, Y_1) \quad (X_2, Y_2)$

Step 1: data collection

$(X_1, Y_1) \Longrightarrow h_1$ (fair classifier)

$(X_2, Y_2) \Longrightarrow h_2$ (fair classifier)

Step 2: offline training and evaluation (separately)

Toy Example: $X_1 \rightarrow Y_1 \rightarrow X_2 \rightarrow Y_2$

(X_1, Y_1) (X_2, Y_2)

Step 1: data collection

$(X_1, Y_1) \Longrightarrow h_1$ (fair classifier)

$(X_2, Y_2) \Longrightarrow h_2$ (fair classifier)

Step 2: offline training and evaluation (separately)

$\hat{X}_1 \xrightarrow{h_1} \hat{Y}_1$ (fair)

$\hat{X}_2 \xrightarrow{h_2} \hat{Y}_2$ (unfair)

Step 3: deploy and make decisions on new data

Toy Example: $X_1 \rightarrow Y_1 \rightarrow X_2 \rightarrow Y_2$

$(X_1, Y_1) \quad (X_2, Y_2)$

Step 1: data collection

$(X_1, Y_1) \Longrightarrow h_1$ (fair classifier)

$(X_2, Y_2) \Longrightarrow h_2$ (fair classifier)

Step 2: offline training and evaluation (separately)

$\hat{X}_1 \xrightarrow{h_1} \hat{Y}_1$ (fair)

$\hat{X}_2 \xrightarrow{h_2} \hat{Y}_2$ (unfair)

Step 3: deploy and make decisions on new data

Why ?

- Decision $\hat{Y}_1$ will affect values of $\hat{X}_2$
- Distribution $X_2 \neq$ Distribution $\hat{X}_2$

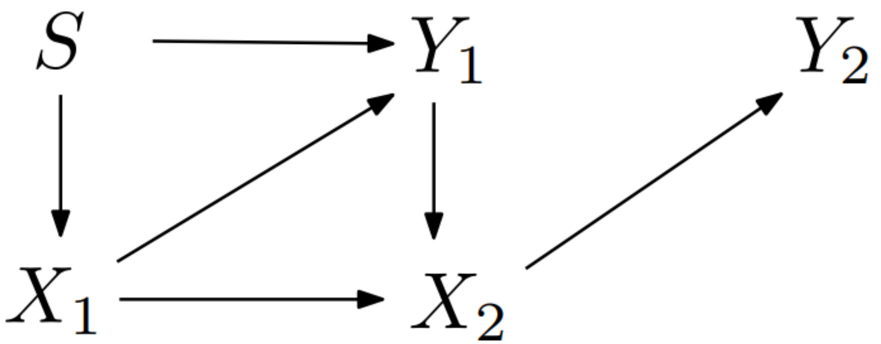
Proposed Solution

- <u>Core idea:</u> leverage Pearl's structural causal model, treat each decision model as a soft intervention, and infer the post-intervention distributions to formulate the loss function as well as the fairness constraints.
- Advantages:
  - Learn multiple fair classifiers simultaneously and only require static training data.
  - Can employ off-the-shelf classification models and optimization algorithms.
  - Achieve causal fairness (total effect in this work).

Using Soft Interventions to Simulate Decision Model Deployments

$Y_1 = h_1(X_1)$

$Y_2 = h_2(X_2) \mid do(Y_1 = h_1(X_1))$



$X_2 \mid do(Y_1 = h_1(X_1))$

- In general, we have l decisions $\{Y_1, \cdots, Y_l\}$.
- For each decision Y_k, we build a classifier $h_k(\boldsymbol{z}_k)$.
- The soft intervention for deploying all these models is $do(h_1, \cdots, h_l)$.

Loss Function and Fair Constraints

- Traditionally, classification error of classifier $h: \boldsymbol{Z} \mapsto Y$ is

$$R(h) = \mathbb{E}_{\boldsymbol{Z}}\left[P(Y = 1|\boldsymbol{z})\mathbf{1}_{h(\boldsymbol{z})<0} + P(Y = 0|\boldsymbol{z})\mathbf{1}_{h(\boldsymbol{z})\geq 0}\right]$$

- Under soft intervention of deploying all models, for classifier h_k

$$R(h_k) = \mathbb{E}_{\boldsymbol{Z}_k|do(h_1,\cdots,h_l)}\left[P(Y_k = 1|\boldsymbol{z}_k)\mathbf{1}_{h(\boldsymbol{z}_k)<0} + P(Y_k = 0|\boldsymbol{z}_k)\mathbf{1}_{h(\boldsymbol{z}_k)\geq 0}\right]$$

- Similarly, fairness constraints is given by total effect

$$TE(h_k) = P(Y_k = 1|do(S = 1, h_1, \cdots, h_l)) - P(Y_k = 1|do(S = 0, h_1, \cdots, h_l))$$

Deriving Loss Function and Fair Constraints with Observed Data

- Loss function

$$R_\phi(h_k) = \mathop{\mathbb{E}}_{S,\mathbf{X}'_{Y_k}}\Bigg[P(y_k^+|\mathbf{z}_k)\phi(h_k(\mathbf{z}_k))\sum_{\mathbf{Y}'_{Y_k}}\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^+}\phi(-h_i(\mathbf{z}_i))\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^-}\phi(h_i(\mathbf{z}_i))\prod_{X_i\in\mathbf{X}'_{Y_k}}\frac{P(\mathbf{y}'_{X_i}|s,x_i,\mathbf{x}'_{X_i})}{P(\mathbf{y}'_{X_i}|s,\mathbf{x}'_{X_i})}$$
$$+ P(y_k^-|\mathbf{z}_k)\phi(-h_k(\mathbf{z}_k))\sum_{\mathbf{Y}'_{Y_k}}\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^+}\phi(-h_i(\mathbf{z}_i))\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^-}\phi(h_i(\mathbf{z}_i))\prod_{X_i\in\mathbf{X}'_{Y_k}}\frac{P(\mathbf{y}'_{X_i}|s,x_i,\mathbf{x}'_{X_i})}{P(\mathbf{y}'_{X_i}|s,\mathbf{x}'_{X_i})}\Bigg].$$

- Fairness constraint

$$T_\phi(h_k) = \mathop{\mathbb{E}}_{\mathbf{X}'_{Y_k}|S=s^+}\Bigg[\phi(-h_k(\mathbf{z}_k))\sum_{\mathbf{Y}'_{Y_k}}\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^+}\phi(-h_i(\mathbf{z}_i))\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^-}\phi(h_i(\mathbf{z}_i))\prod_{X_i\in\mathbf{X}}\frac{P(\mathbf{y}'_{X_i}|s^+,x_i,\mathbf{x}'_{X_i})}{P(\mathbf{y}'_{X_i}|s^+,\mathbf{x}'_{X_i})}\Bigg]$$
$$+ \mathop{\mathbb{E}}_{\mathbf{X}'_{Y_k}|S=s^-}\Bigg[\phi(h_k(\mathbf{z}_k))\sum_{\mathbf{Y}'_{Y_k}}\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^+}\phi(-h_i(\mathbf{z}_i))\prod_{Y_i\in\mathbf{Y}'_{Y_k},y_i^-}\phi(h_i(\mathbf{z}_i))\prod_{X_i\in\mathbf{X}}\frac{P(\mathbf{y}'_{X_i}|s^-,x_i,\mathbf{x}'_{X_i})}{P(\mathbf{y}'_{X_i}|s^-,\mathbf{x}'_{X_i})}\Bigg] - 1.$$

Problem Formulation for Fair Multiple Decision Making

- *The problem of fair multiple decision making for* $\boldsymbol{Y} = \{Y_1, \cdots, Y_l\}$ *is formulated as the following constrained optimization problem:*

$$\min_{h_1,\cdots,h_l \in \mathcal{H}} \sum_{k=1}^{l} R_\phi(h_k) \quad s.t. \quad \forall k, -\tau_k \le T_\phi(h_k) \le \tau_k$$

where $R_\phi(h_k)$ *and* $T_\phi(h_k)$ *are loss function and fair constraint.*

- Can be solved easily using gradient-based algorithms as each h_k is involved as a single term in the multiplication.

Excess Risk Bound

- *For any classification-calibrated surrogate function* ϕ *satisfying* $\phi(0) = 1$ *and* $\inf_{\alpha \in \mathbb{R}} \phi(\alpha) = 0$, *any measurable function* h_k *for predicting* Y_k, *we have*

$$\psi(R(h_k) - R^*) \leq R_\phi(h_k) - R_\phi^*$$

where ψ *is a non-decreasing function mapping from* $[0,1]$ *to* $[0, \infty)$.

Experiments

- Data:
  - Synthetic data: Manually define a causal graph and conditional probability tables. Data is generated by sampling each attribute in topological order according to the conditional probability.
  - Adult: Build the causal graph by using the PC algorithm. Age is treated as S, Workclass and Income are treated as Y_1 and Y_2.
- Baselines:
  - Separate method: Each classifier is learned separately on training data.
  - Serial method: Classifiers are learned sequentially following the topological order of the causal graph.
- Our method (joint method): formulates the optimization problem on the training data to learn all classifiers simultaneously

UNIVERSITY OF ARKANSAS

Experiments

Table 1: Accuracy and unfairness from Unconstrained, Separate, Serial and Joint methods on synthetic and Adult data (bold values indicate violation of fairness).

Phase			Synthetic				Adult			
			Uncons.	Separate	Serial	Joint	Uncons.	Separate	Serial	Joint
Train	h_1	Acc. (%)	80.32	75.35	75.35	75.35	55.71	55.64	55.63	55.63
		Unfairness	**0.15**	0.01	0.01	0.01	**0.15**	0.05	0.05	0.05
	h_2	Acc. (%)	90.13	75.79	84.02	82.77	76.75	71.17	68.90	69.31
		Unfairness	**0.23**	0.04	0.03	0.04	**0.24**	0.10	0.10	0.10
Test	h_1	Acc. (%)	80.70	75.54	75.54	75.54	55.63	55.56	55.57	55.57
		Unfairness	**0.15**	0.01	0.01	0.01	**0.15**	0.05	0.05	0.05
	h_2	Acc. (%)	89.95	77.06	84.16	82.09	77.07	73.33	68.91	69.40
		Unfairness	**0.13**	**0.09**	0.03	0.03	**0.23**	**0.17**	0.10	0.10

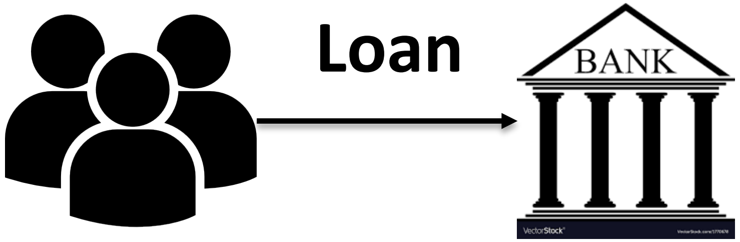
FAIR REPEATED DECISION MAKING WITH LONG-TERM IMPACTS

Y. Hu and L. Zhang. Achieving Long-term Fairness in Sequential Decision Making. AAAI, 2022.

Sequential Decisions

- In practice, decision making systems are usually operating in a dynamic manner such that the classifier makes sequential decisions over a period of time.

<u>Example</u>:



The bank uses the applicants' credit scores to make loan decisions

The bank's decisions will affect the applicants' credit scores

Feedback Loop

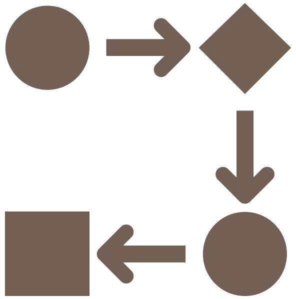
Long-term Fairness

- Fair decision making should concern not only the fairness of a single decision but more importantly, whether a decision model can impose fair long-term effects on different groups.

- This notion of fairness is referred to as long-term fairness in recent studies.

Challenges

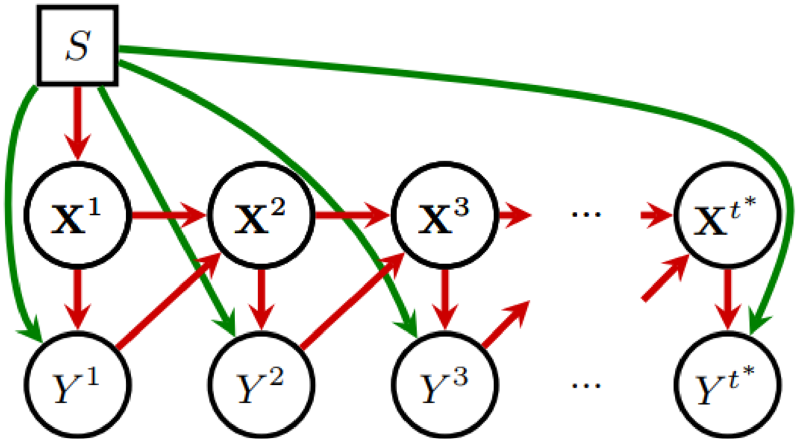
Define notions and quantitative measures for long-term fairness.



Develop efficient learning algorithms for dynamic systems with repeated feedback loops.

Causality-based Long-term Fairness

- Based on SCM, we assume a time-lagged causal graph G for describing the causal relationship among variables over time.



Soft intervention:

Causality-based Long-term Fairness

- **long-term fairness.** Formulated as path-specific effects that are transmitted in the time-lagged causal graph along certain paths.

$$\boldsymbol{X} \begin{cases} \boldsymbol{X}_i & \text{irrelevant attributes: justifiable in decision making, evolved autonomously or altered by external factors.} \\ \boldsymbol{X}_r & \text{relevant attributes: the remaining} \end{cases}$$

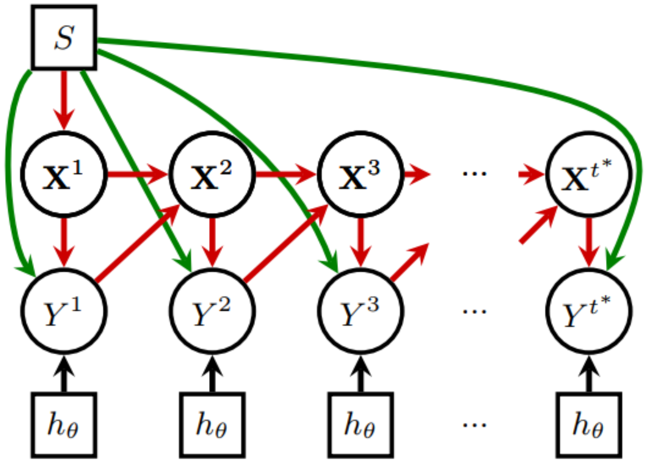
Definition 1 (Long-term Fairness). *The long-term fairness of a decision model* h_θ is measured by $P\left(\hat{Y}^{t*}(s_\pi^+, \theta)\right) - P\left(\hat{Y}^{t*}(s_\pi^-, \theta)\right)$ where π is a set of paths from S to $\hat{Y}^{t*}$ passing through $\boldsymbol{X}_r^1, \hat{Y}^1, \ldots, \boldsymbol{X}_r^{t*-1}, \hat{Y}^{t*-1}, \boldsymbol{X}_r^{t*}$, s_π represents the path-specific hard intervention and θ represents the soft intervention through all paths.

Loss Function and Short-term Fairness

- Two other requirements:

    - **Short-term Fairness.** The decision model should also satisfy certain short-term fairness requirement at each time step to enforce local equality, which may be stipulated by law or regulations.

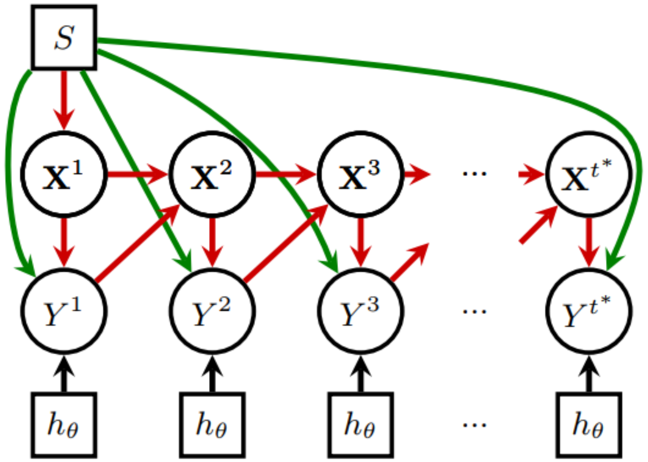
Definition 2 (Short-term Fairness). *The short-term fairness of a decision model* h_θ at time t is measured by the causal effect transmitted through paths involved in time t , i.e., $P\left(\hat{Y}^t(s^+_{\pi^t},\theta)\right) - P\left(\hat{Y}^t(s^-_{\pi^t},\theta)\right)$, where $\pi^t = \{S \to \widetilde{\boldsymbol{X}}_r \to \hat{Y}^t, S \to \hat{Y}^t\}$ with redlining attributes $\widetilde{\boldsymbol{X}}_r$, s_π is the path-specific hard intervention and θ represents the soft intervention.



Loss Function and Short-term Fairness

- Two other requirements:
  - **Institution Utility.** It is a natural desire for a predictive decision model to maximize the institution utility.

Definition 3 (Institute Utility). *The institution utility of a decision model* h_θ is measured by the aggregate loss given by $\sum_{t=1}^{t*} E\left[L\left(Y^t, \hat{Y}^t\right)\right]$ where $L(\cdot)$ is the loss function.



Soft Intervention for Model Deployment

- In all three definitions, we use soft intervention for modeling the decision model deployment.
  - We treat the deployment of the decision model at each time step as to perform a soft intervention on the decision variable.
  - The change to underlying population could be inferred as the post-intervention distribution after performing the soft intervention.

Learning Fair Decision Model

The goal is to learn a functional mapping $h_\theta\colon (S, \boldsymbol{X}^t) \rightarrow Y^t$ parameterized with θ. By meeting the two requirements of institution utility and short-term fairness, the functional mapping will achieve long-term fairness.

Problem Formulation 1. *The problem of fair sequential decision making is formulated as the constrained optimization:*

$$\arg\min_{\theta} \sum_{t=1}^{t*} E\left[L\left(Y^t, \hat{Y}^t\right)\right]$$

$$s.t.\ \ P\left(\hat{Y}^{t*}(s_\pi^+, \theta) = 1\right) - P\left(\hat{Y}^{t*}(s_\pi^-, \theta) = 1\right) \leq \tau_l$$

$$P\left(\hat{Y}^t\left(s_{\pi^t}^+, \theta\right) = 1\right) - P\left(\hat{Y}^t\left(s_{\pi^t}^-, \theta\right) = 1\right) \leq \tau_t,\ t = 1, \ldots, t^*$$

where τ_l and τ_t are thresholds for long-term fairness and short-term fairness constraints, respectively.

Performative Risk Optimization

- Solving the optimization problem in Problem Formulation 1 is not trivial.
- Convert Problem Formulation 1 to the form of performative risk optimization.
- The general formulation of the performative risk optimization is

$$\mathrm{PR}(\theta) = \mathop{\mathbb{E}}_{Z \sim \mathcal{D}(\theta)} \ell(Z;\theta)\,.$$

Performative Risk Optimization

- Reformulate utility, short-term fairness and long-term fairness in the form of performative risk.

$$l_u(\theta) = \sum_{t=1}^{t^*} \mathop{\mathbb{E}}_{S,\mathbf{X}^t,Y^t \sim P(S,\mathbf{X}^t,Y^t)} \left[\phi\left(Y^t h_\theta(\mathbf{X}^t, S)\right)\right],$$

$$l_l(\theta) = \frac{1}{2}\left\{ \mathop{\mathbb{E}}_{\mathbf{X}^{t^*} \sim P(\mathbf{X}^{t^*}(s_\pi^+,\theta))} \left[\phi\left(-h_\theta\left(\mathbf{X}^{t^*}, s^-\right)\right)\right] \right.$$
$$\left. + \mathop{\mathbb{E}}_{\mathbf{X}^{t^*} \sim P(\mathbf{X}^{t^*}(s_\pi^-,\theta))} \left[\phi\left(h_\theta\left(\mathbf{X}^{t^*}, s^-\right)\right)\right] - 1 - \tau_l \right\},$$

$$l_s(\theta) = \frac{1}{t^*} \sum_{t=1}^{t^*} \left\{ \mathop{\mathbb{E}}_{\mathbf{X}^t \sim P(\mathbf{X}^t(s_{\pi^t}^+,\theta))} \left[\phi\left(-h_\theta\left(\mathbf{X}^t, s^-\right)\right)\right] \right.$$
$$\left. + \mathop{\mathbb{E}}_{\mathbf{X}^t \sim P(\mathbf{X}^t(s_{\pi^t}^-,\theta))} \left[\phi\left(h_\theta\left(\mathbf{X}^t, s^-\right)\right)\right] - 1 - \tau_t \right\}.$$

Performative Risk Optimization

Problem Formulation 2. *The problem of fair sequential decision making is reformulated as the performative risk optimization:*

$$\arg\min_{\theta} l(\theta) = \lambda_u l_u(\theta) + \lambda_l l_l(\theta) + \lambda_s l_s(\theta)$$

where λ_u, λ_l *and* λ_s *are weight parameters and satisfy* $\lambda_u + \lambda_l + \lambda_s = 1$.

Repeated Risk Minimization

- Repeated risk minimization (RRM) is an iterative algorithmic heuristic for solving the performative risk optimization problem.

Algorithm 1: Repeated Risk Minimization

Input : Dataset $\mathcal{D} = \{(S, \mathbf{X}^t, Y^t)\}_{t=1}^{l}$, time -lagged causal graph $\mathcal{G}$, convergence threshold δ

Output: The stable model h_θ

Train a classifier on $\mathcal{D}$ according to Eq. (2) without the soft intervention to obtain the initial parameter θ_0;
$i \leftarrow 0$;
repeat
    Sampled the post-intervention distributions $P\left(\mathbf{X}^{t^*}(s_\pi^+, \theta_i)\right)$ and $P\left(\mathbf{X}^{t^*}(s_\pi^-, \theta_i)\right)$;
    Sampled the post-intervention distributions $P\left(\mathbf{X}^t(s_\pi^+, \theta_i)\right)$ and $P\left(\mathbf{X}^t(s_\pi^-, \theta_i)\right)$ for each t;
    Minimize $l(\theta)$ according to Eq. (2) to obtain θ_{i+1};
    $\triangle = \|\theta_{i+1} - \theta_i\|_2$;
    $i \rightarrow i + 1$;
until $\triangle < \delta$;
$\theta \leftarrow \theta_i$;
return h_θ;

Convergence Analysis of RRM

- The convergence of the RRM algorithm depends on the smoothness and convexity of the loss function, as well as the sensitivity of the distribution to the parameters.

Theorem 1. *Suppose that surrogated loss function* $(\varphi \circ h)(\cdot)$ *is* β *–jointly smooth and* γ*-strongly convex, and suppose that* $\boldsymbol{X}^{t+1}$ *are* c*-sensitive for any* t*, then the repeated risk minimization converges to a stable point at a linear rate, if* $2mc(t^{*}-1) < \frac{\beta}{\gamma}$.

Experiments

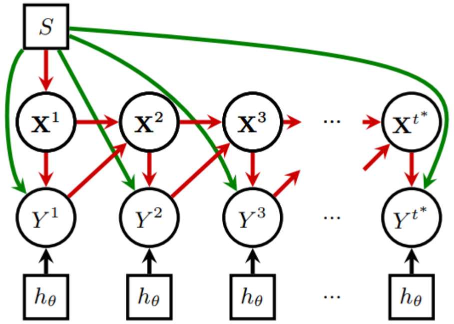
- Baselines:

  - **Logistic Regression (LR)**: An unconstrained logistic regression model which takes user features and labels from all time steps as inputs and outputs.

  - **Fair Model with Demographic Parity (FMDP)**: On the basis of the logistic regression model, fairness constraint is added to achieve demographic parity.

  - **Fair Model with Equal Opportunity (FMEO)**: On the basis of the logistic regression model, fairness constraint is added to achieve equal opportunity.

Experiments



- Synthetic Data:

We simulate a process of bank loans following the above time-lagged causal graph, where S is the protected attribute like race, $\boldsymbol{X}^t$ represents the financial status of applicants, and Y^t represents the decisions about whether to grant loans.

We sample the predicted decisions from:

$$P(\hat{Y}^t) = \sigma(h_{\theta^*}(\mathbf{X}^t, S)),\ \hat{Y}^t \sim 2 \cdot \text{Bernoulli}(P(\hat{Y}^t)) - 1.$$

$\boldsymbol{X}^{t+1}$ is generated according to the update rule below:

$$\mathbf{X}^{t+1} = \begin{cases} \mathbf{X}^t - \epsilon \cdot \theta^t + b & \hat{Y}^t = 1, Y^t = -1 \\ \mathbf{X}^t + \epsilon \cdot \theta^t + b & \hat{Y}^t = 1, Y^t = 1 \\ \mathbf{X}^t + b & \hat{Y}^t = -1 \end{cases}$$

Experiments

- Semi-synthetic Data:

- Use the Taiwan credit card dataset (Yeh and Lien 2009) as the initial data at t = 1
- Extract 3000 samples and choose two features PAY AMT1 and PAY AMT2
- Generate a 4-step dataset using similar update rule

Yeh, I.-C.; and Lien, C.-h. 2009. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. Expert Systems with Applications, 36(2): 2473–2480.

Experiments

- Results of Synthetic Data:

Table 1: Accuracy, short-term and long-term fairness of different algorithms on the synthetic dataset.

Alg.	Metric	Time steps				
		$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
RL	Acc	0.912	0.894	0.917	0.921	0.917
	Short	0.152	0.160	0.166	0.164	0.174
	Long	0.058	0.117	0.173	0.246	0.340
FMDP	Acc	0.735	0.706	0.704	0.708	0.725
	Short	0.212	0.216	0.224	0.220	0.232
	Long	0.180	0.306	0.376	0.431	0.481
FMEO	Acc	0.829	0.790	0.795	0.800	0.814
	Short	0.010	0.010	0.010	0.014	0.020
	Long	0.080	0.122	0.190	0.276	0.352
Ours	Acc	0.801	0.754	0.729	0.707	0.692
	Short	0.012	0.008	0.012	0.008	0.002
	Long	0.040	0.024	0.020	0.012	0.002

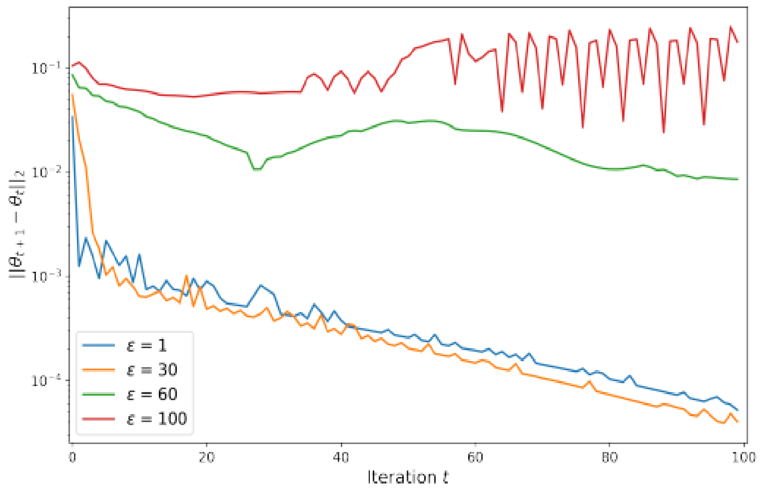


Figure 2: The convergence results for different values of ϵ on the synthetic dataset.

Experiments

- Results of Semi-synthetic Data:

Table 2: Accuracy, short-term and long-term fairness of different algorithms on the semi-synthetic dataset.

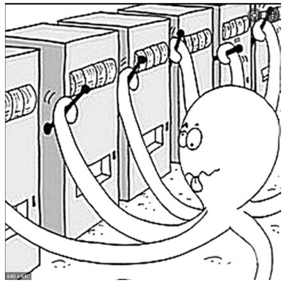
Alg.	Metric	Time steps			
		$t=1$	$t=2$	$t=3$	$t=4$
RL	Acc	0.828	0.826	0.841	0.816
	Short	0.015	0.018	0.021	0.012
	Long	0.038	0.088	0.243	0.433
FMDP	Acc	0.830	0.843	0.846	0.841
	Short	0.063	0.066	0.075	0.069
	Long	0.038	0.076	0.223	0.397
FMEO	Acc	0.824	0.830	0.830	0.813
	Short	0.072	0.075	0.087	0.078
	Long	0.006	0.045	0.156	0.295
Ours	Acc	0.648	0.648	0.680	0.687
	Short	0.006	0.006	0.003	0.006
	Long	0.064	0.043	0.016	0.003

FAIR RECOMMENDATION IN ONLINE STOCHASTIC SETTINGS

W. Huang, L. Zhang, and X. Wu. Achieving Counterfactual Fairness for Causal Bandit. AAAI, 2022.

Bandit Recommendation

❖ Bandit Algorithm



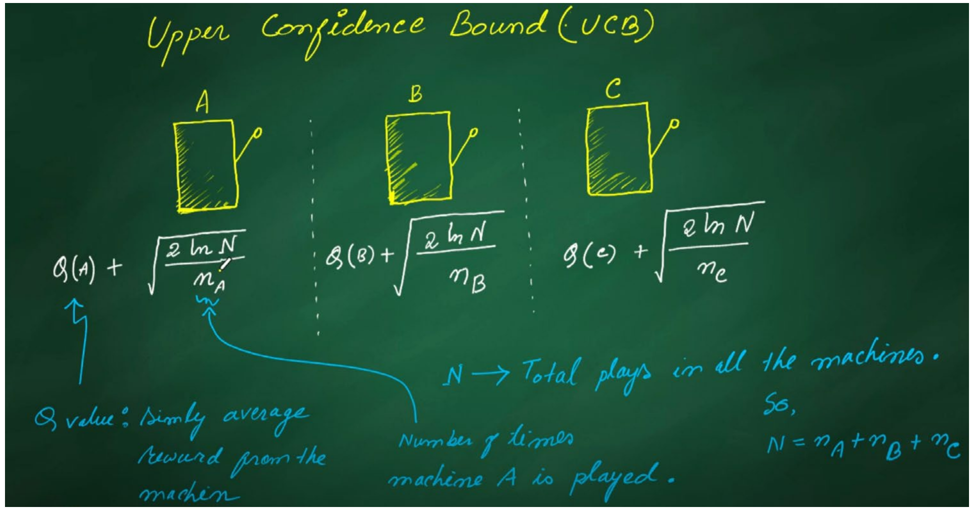
- Time horizon: $t = 1,2, \ldots, T$
- At each time step:
  - Pull an arm a_t
  - Receive a reward r_t

Goal: Maximize $\sum_{t=1}^{T} r_t$

Policy: Implies which arm to pull at each round.

Evaluation metric: Cumulative regret $\mathcal{R}_T = \sum_{t=1}^{T} (\mu_{a^*} - \mu_{a_t})$

❖ UCB Method



- At each round, updates the upper confidence bound (UCB) of the reward for each arm.
- Policy: Pull the arm with the largest UCB at each round.

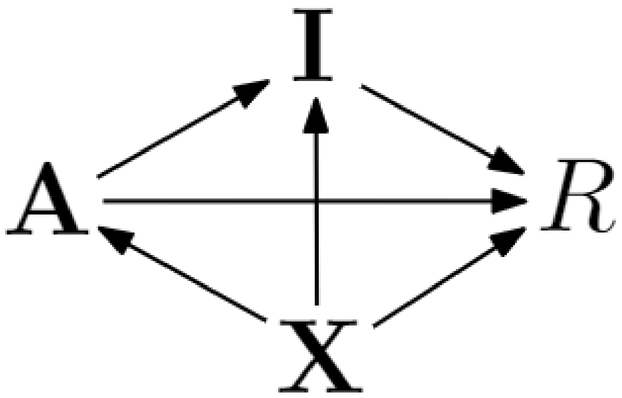
Counterfactual Fairness in Bandits

❖ Counterfactual Fairness

Will customers with similar profiles receive similar rewards regardless of their protected attributes and recommended items?

Goal: Achieve user-side individual fairness for customers in bandit online recommendation.

Causal Bandits



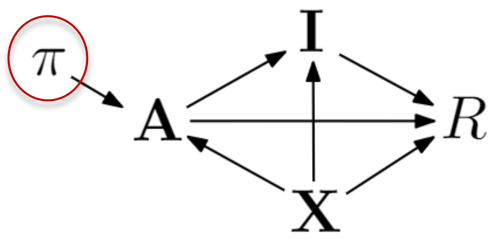
$\boldsymbol{A}$: Arm features

$\mathbf{X}$: User features

R : Reward

$\mathbf{I}$: Intermediate features between $\boldsymbol{A}$ and R

Modeling Arm Selection via Soft Intervention



$\boldsymbol{A}$ **:** Arm features
$\mathbf{X}$ **:** User features
R **:** Reward
$\mathbf{I}$ **:** Intermediate features between $\boldsymbol{A}$ and R
π **:** Soft intervention conducted on arm selection

Definition of π

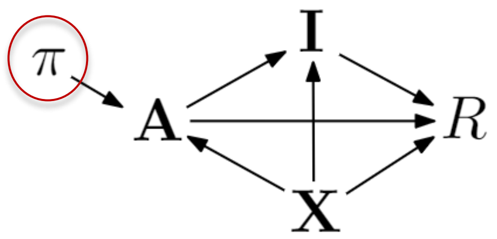
- An **arm selection strategy** conducted on $\boldsymbol{A}$ while user features $\mathbf{X}$ and all other relationships in the causal graph are **unchanged**.

Advantages

- Able to capture the complex causal relationship.
- Flexible in terms of the functional form.
- Either deterministic or stochastic.

Strong generalization property to depict various bandit algorithms by adopting different soft interventions.

Modeling Arm Selection via Soft Intervention



$\boldsymbol{A}$ **:** Arm features
$\mathbf{X}$ **:** User features
R **:** Reward
$\mathbf{I}$ **:** Intermediate features between $\boldsymbol{A}$ and R
π **:** Soft intervention conducted on arm selection

- Expected reward under soft intervention π

$$\mu_\pi = \mathbb{E}[R(\pi)|\mathbf{x}_t] = \overbrace{\sum_{\mathbf{a}} \underbrace{P_\pi(\mathbf{a}|\mathbf{x}_t)}_{\text{Distribution defined by policy } \pi} \cdot \mathbb{E}[R(\mathbf{a})|\mathbf{x}_t]}^{\text{Decomposition based on soft intervention}} = \mathbb{E}_{\mathbf{a}\sim\pi}\left[\mathbb{E}[R(\mathbf{a})|\mathbf{x}_t]\right]$$

Counterfactual Fairness in bandit setting: Definition and Bound

How to measure individual level user-side fairness in terms of the reward?

- A protected attribute: $S \in \mathbf{X}$
- The counterfactual reward by setting $S = s^*$: $\mathbb{E}[R(\pi, s^*)|\mathbf{x}_t]$
- The counterfactual discrepancy regarding to a policy

$$\Delta_\pi = \mathbb{E}[R(\pi, s^+)|\mathbf{x}_t] - \mathbb{E}[R(\pi, s^-)|\mathbf{x}_t]$$

Counterfactual Fairness in bandit setting: Definition and Bound

How to measure individual level user-side fairness in terms of the reward?

- Definition of a τ-counterfactually fair policy

A policy π is counterfactually fair for an individual arrived if $\Delta_\pi = 0$. The policy is τ- counterfactually fair if $|\Delta_\pi| \leq \tau$ where τ is the predefined fairness threshold.

- Bound under the unidentifiable case

If there exists a non-empty set $\mathbf{B} \subseteq \mathbf{X} \backslash \{S\}$ which are descendants of S, then $\mu_{a,s^*} = \mathbb{E}[R(a, s^*)|\mathbf{x}_t]$ is bounded by

$$\sum_{\mathbf{z}} \min_{\mathbf{b}} \{\mathbb{E}[R|s^*, \mathbf{w} \backslash s_t]\} \cdot P(\mathbf{z}|\mathbf{x}_{t,a}) \leq \mu_{a,s^*} \leq \sum_{\mathbf{z}} \max_{\mathbf{b}} \{\mathbb{E}[R|s^*, \mathbf{w} \backslash s_t]\} \cdot P(\mathbf{z}|\mathbf{x}_{t,a})$$

F-UCB: A Counterfactually Fair Causal Bandit

Goal: Achieving counterfactual fairness for causal bandit.

Main idea: Select optimal policy in a **counterfactually safe** policy subspace at each round.

- Estimated reward mean of a policy $$\hat{\mu}_{\pi}(t) = \mathbb{E}_{\mathbf{a}\sim\pi}\left[\sum_{\mathbf{Z}} \hat{\mu}_{\mathbf{w}}(t) \cdot P(\mathbf{z}|\mathbf{x}_{t,a})\right] \quad (1)$$
- Estimated counterfactual reward $$\hat{\mu}_{a,s^*}(t) = \sum_{\mathbf{Z}} \hat{\mu}_{\mathbf{w}^*}(t) \cdot P(\mathbf{z}|s^*, \mathbf{x}_{t,a} \backslash s_t)$$
- Estimated fairness discrepancy $$\hat{\Delta}_{\pi}(t) = \left|\mathbb{E}_{\mathbf{a}\sim\pi}[\hat{\mu}_{a,s^+}(t)] - \mathbb{E}_{\mathbf{a}\sim\pi}[\hat{\mu}_{a,s^-}(t)]\right| \quad (2)$$

Algorithm 2 F-UCB: Fair Causal Bandit

```
Input: Policy space Π, fairness threshold τ, confidence level parameter δ, original
causal Graph 𝒢 with domain knowledge
Find the d-separation set W with minimum subset Z in terms of domain space.
for t = 1, 2, 3, ..., T do
    for π ∈ Π_t do
        Compute the estimated reward mean using Equation 1 and the estimated
        fairness discrepancy using Equation 2.
    Determine the conservative fair policy subspace Φ̄_t.
    Find the optimal policy following Equation 3 within Φ̄_t.
    Take action a_t ∼ π_t and observe a real-valued payoff r_t and a d-separation set
    value w_t.
    Update μ̂_w(t) for all w ∈ W.
```

F-UCB: A Counterfactually Fair Causal Bandit

Goal: Achieving counterfactual fairness for causal bandit.

Main idea: Select optimal policy in a **counterfactually safe** policy subspace at each round.

- Construction of the safe policy space: $\bar{\Phi}_t = \{\pi : UCB_{\Delta_\pi}(t) \leq \tau\}$

  where $UCB_{\Delta_\pi}(t) = \hat{\Delta}_\pi(t) + \sum_{\mathbf{Z}} \sqrt{\frac{8\log(1/\delta)}{1 \vee N_{\mathbf{w}}(t)}} P(\mathbf{z}|\mathbf{x}_{t,a})$

- Policy taken at each round:

$$\pi_t = \arg\max_{\pi \in \Pi_t} \mathbb{E}_{\mathbf{a}\sim\pi}[UCB_a(t)] \qquad (3)$$

Algorithm 2 F-UCB: Fair Causal Bandit

1: Input: Policy space Π, fairness threshold τ, confidence level parameter δ, original causal Graph $\mathcal{G}$ with domain knowledge
2: Find the d-separation set $\mathbf{W}$ with minimum subset $\mathbf{Z}$ in terms of domain space.
3: **for** $t = 1, 2, 3, ..., T$ **do**
4:   **for** $\pi \in \Pi_t$ **do**
5:     Compute the estimated reward mean using Equation 1 and the estimated fairness discrepancy using Equation 2.
6:   Determine the conservative fair policy subspace $\bar{\Phi}_t$.
7:   Find the optimal policy following Equation 3 within $\bar{\Phi}_t$.
8:   Take action $\mathbf{a}_t \sim \pi_t$ and observe a real-valued payoff r_t and a d-separation set value $\mathbf{w}_t$.
9:   Update $\hat{\mu}_{\mathbf{w}}(t)$ for all $\mathbf{w} \in \mathbf{W}$.

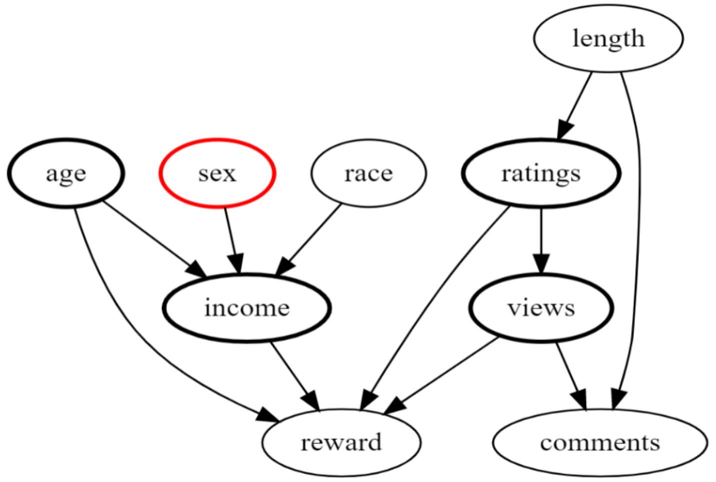
Experiment Results on Adult-YouTube Video Dataset

- **Adult Dataset**

- 31,561 instances: 21,790 males and 10,771 females.
- Each with 8 categorical variables and 3 continuous variables.

- **Youtube video Dataset**

- 1,580 instances.
- Each having 6 categorical features.



τ	Regret	Unfair Decisions	
	F-UCB	F-UCB	Fair-LinUCB
0.1	361.43	0	2053
0.2	332.10	0	1221
0.3	323.12	0	602
0.4	303.32	0	82
0.5	296.19	0	6

Regret for Fair-LinUCB ≈ 250

Huang, W.; Labille, K.; Wu, X.; Lee, D.; and Heffernan, N.2020. Achieving User-Side Fairness in Contextual Bandits. Big Data 2021.

Conclusions and Future Directions

- Explore fair sequential decision making on two dimensions: decision models are deployed sequentially and data arrive sequentially.
- Show that Structural Causal Model can be used as a fair ML framework, where soft intervention is adopted as a general methodology for modeling model deployment.
- Crosspoint of two dimensions: fair multi-stage recommendation where customers are recommended a sequence of items from different stages regarding different but related topics.

Thank you

Lu Zhang
University of Arkansas
lz006@uark.edu

This work was supported in part by NSF.