Mind the GAP: Improving Robustness to Subpopulation Shifts with Group-Aware Priors

Tim G. J. Rudner
@timrudner

Ya Shi Zhang
@andrew_yashi

Andrew Gordon Wilson
@andrewgwils

Julia Kempe
@KempeLab

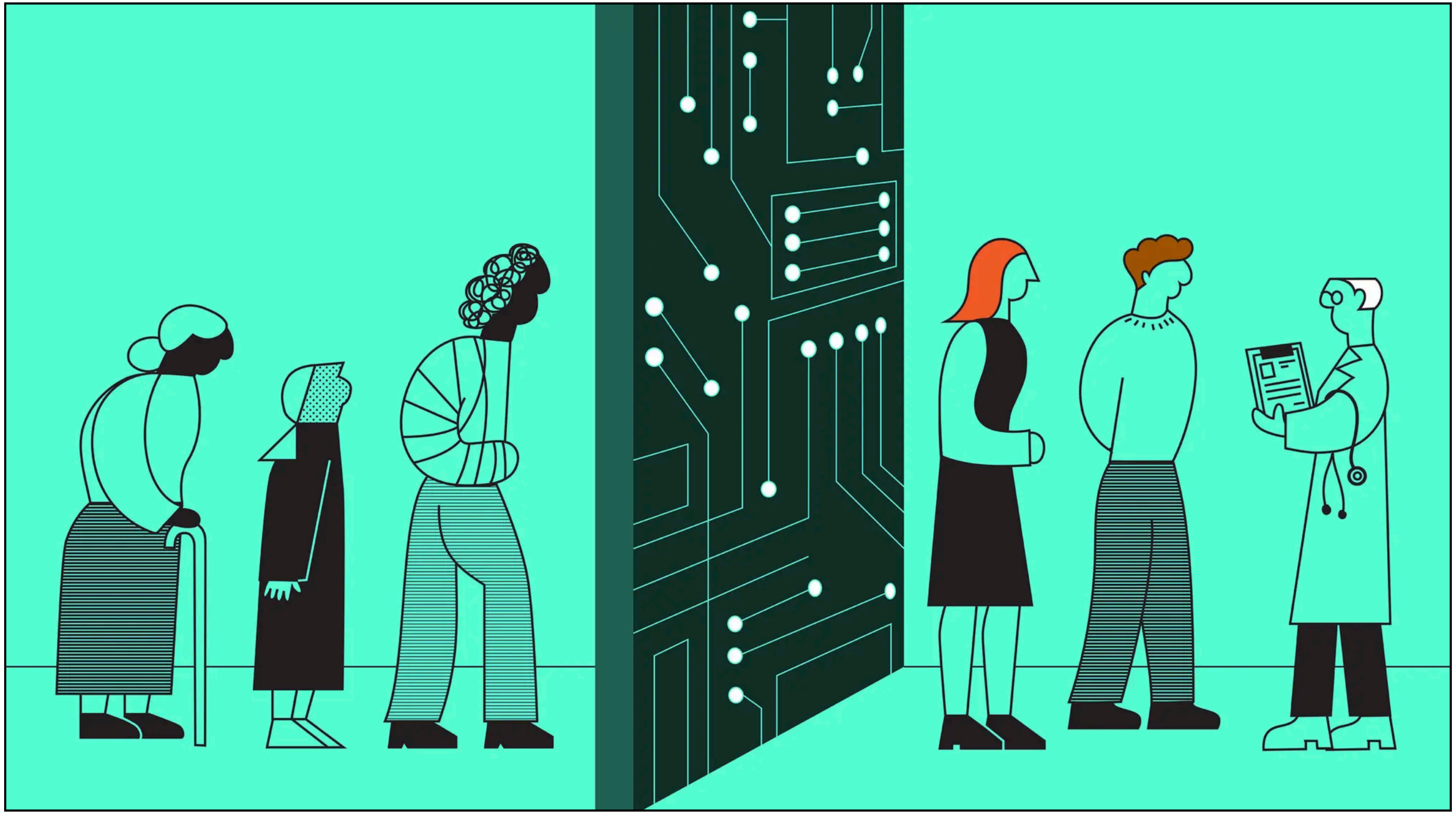
International Conference on Artificial Intelligence and Statistics 2024

NYU

Correspondence to
tim.rudner@nyu.edu

Paper:
timrudner.com/gap

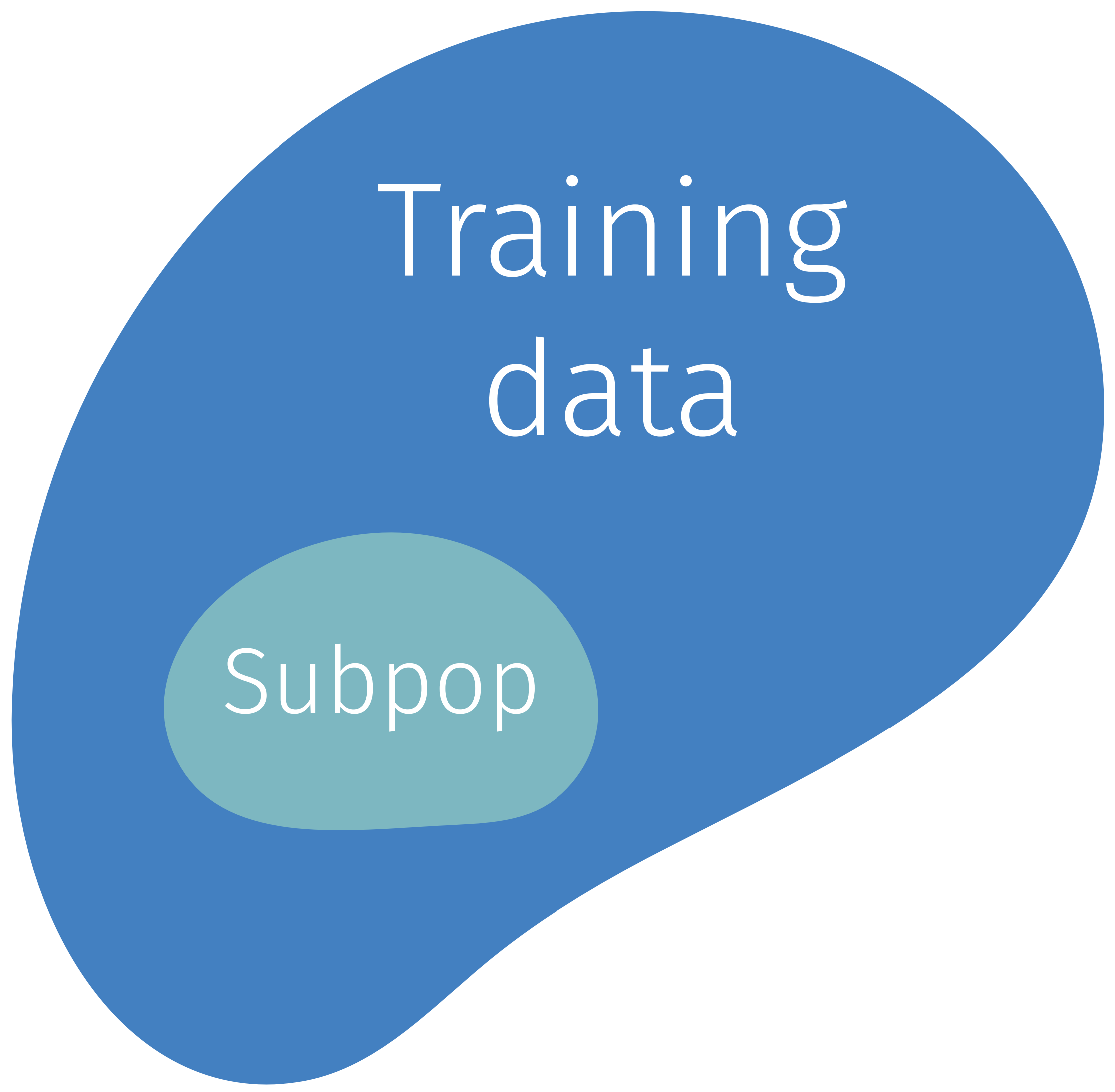
AI can exacerbate existing disparities



Dhruv Khullar. *A.I. Could Worsen Health Disparities*, The New York Times, 2019.

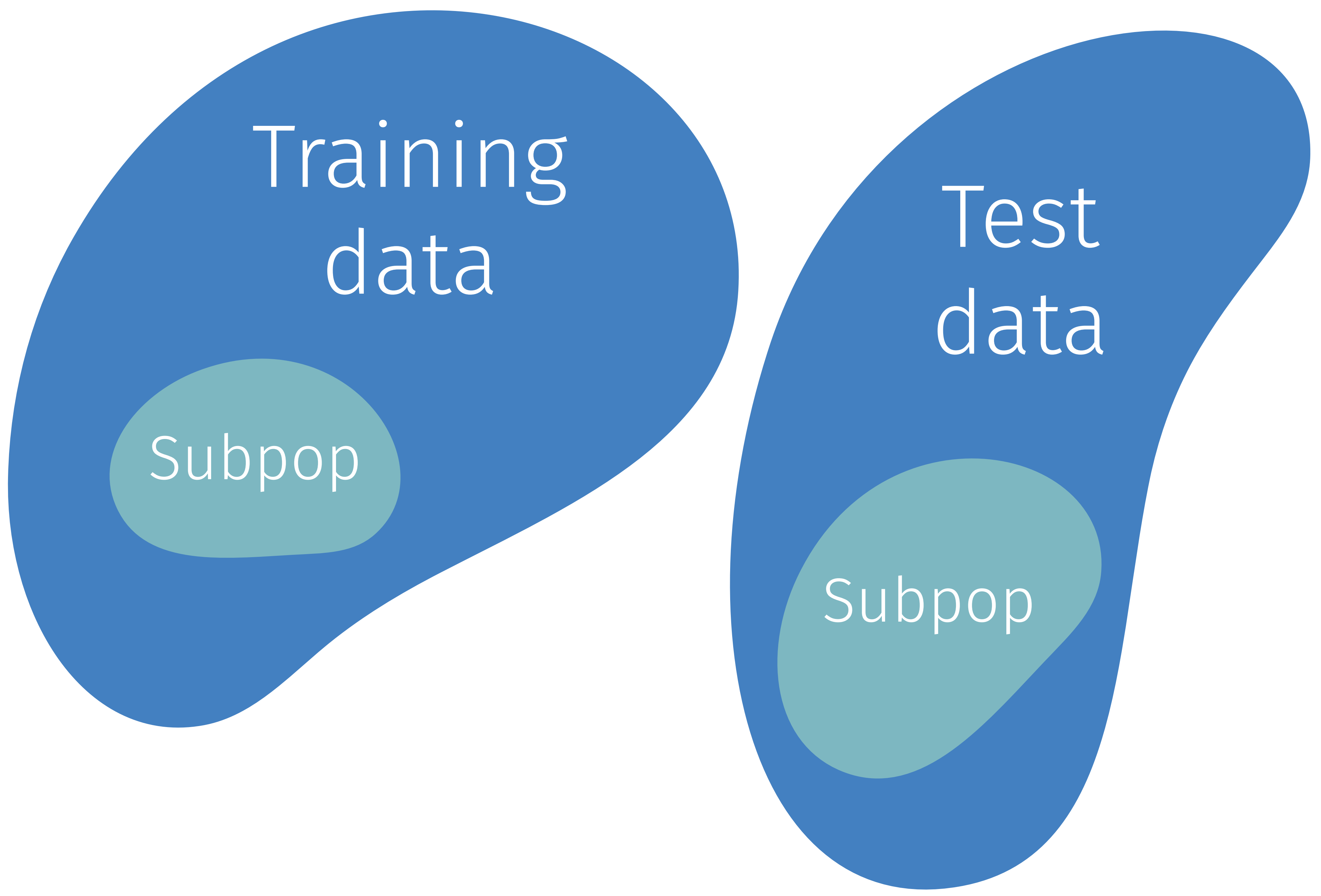
Fair machine learning

We want models that are robust across groups.



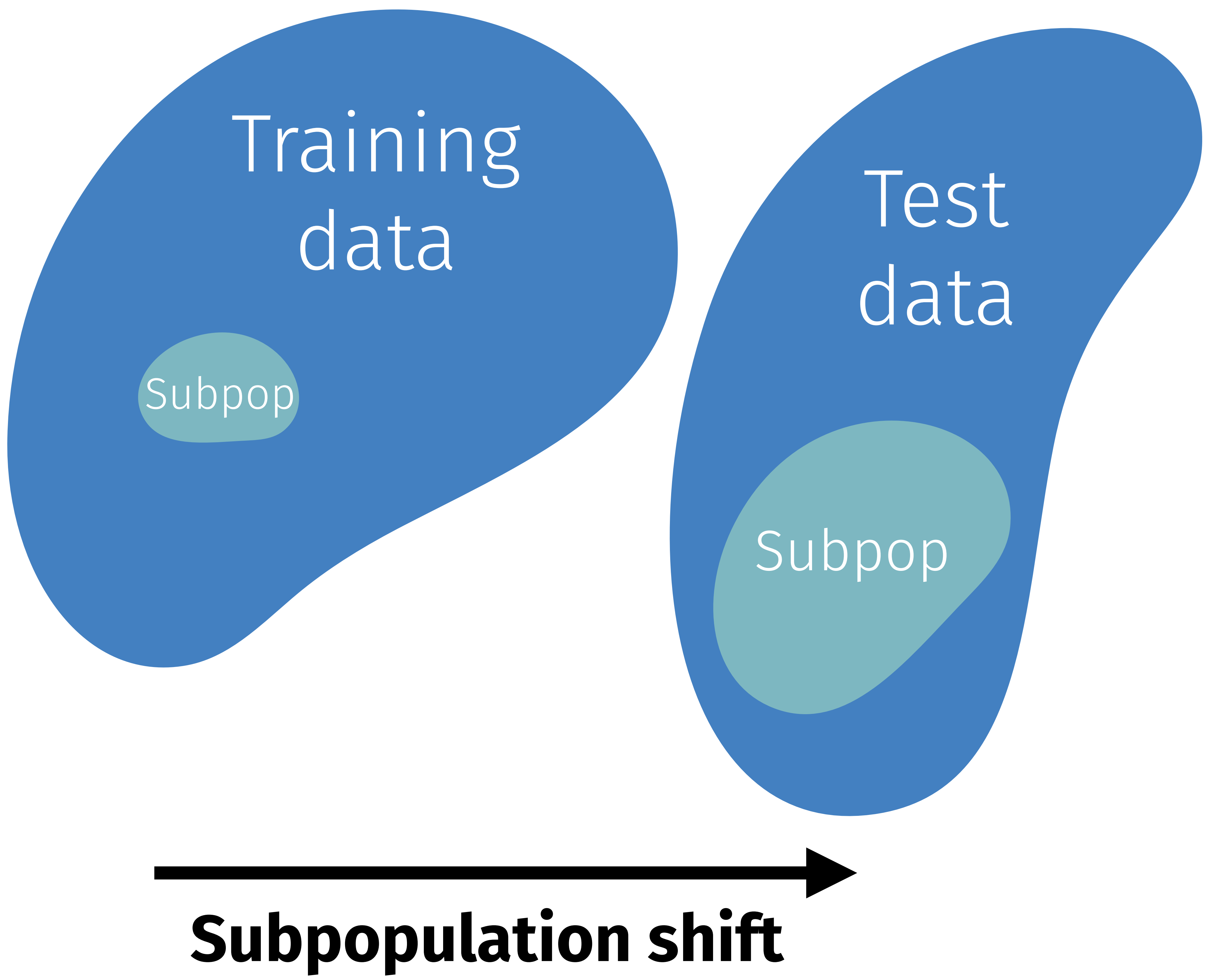
Fair machine learning

We want models that are robust across groups.



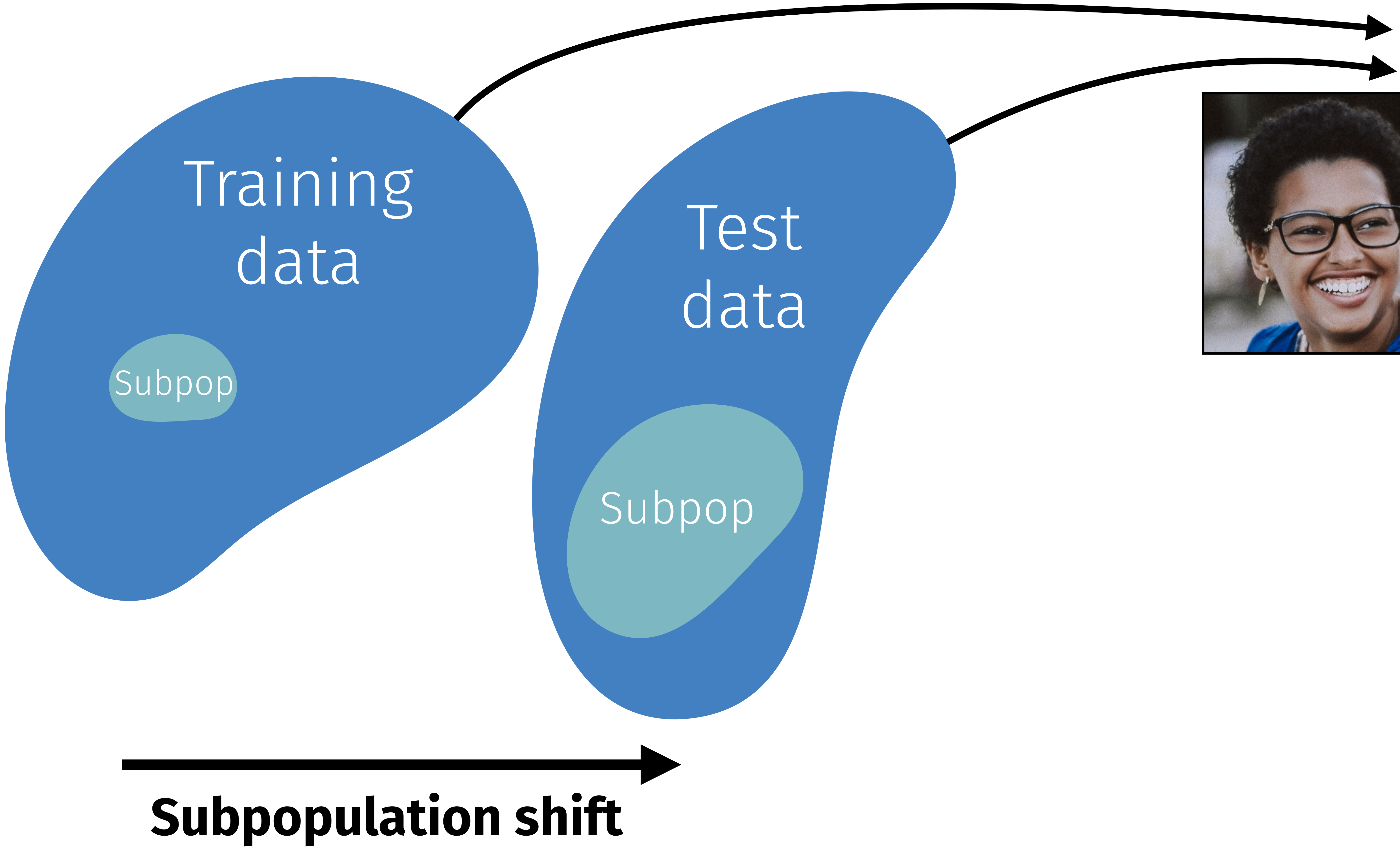
Fair machine learning

We want models that are robust across groups.



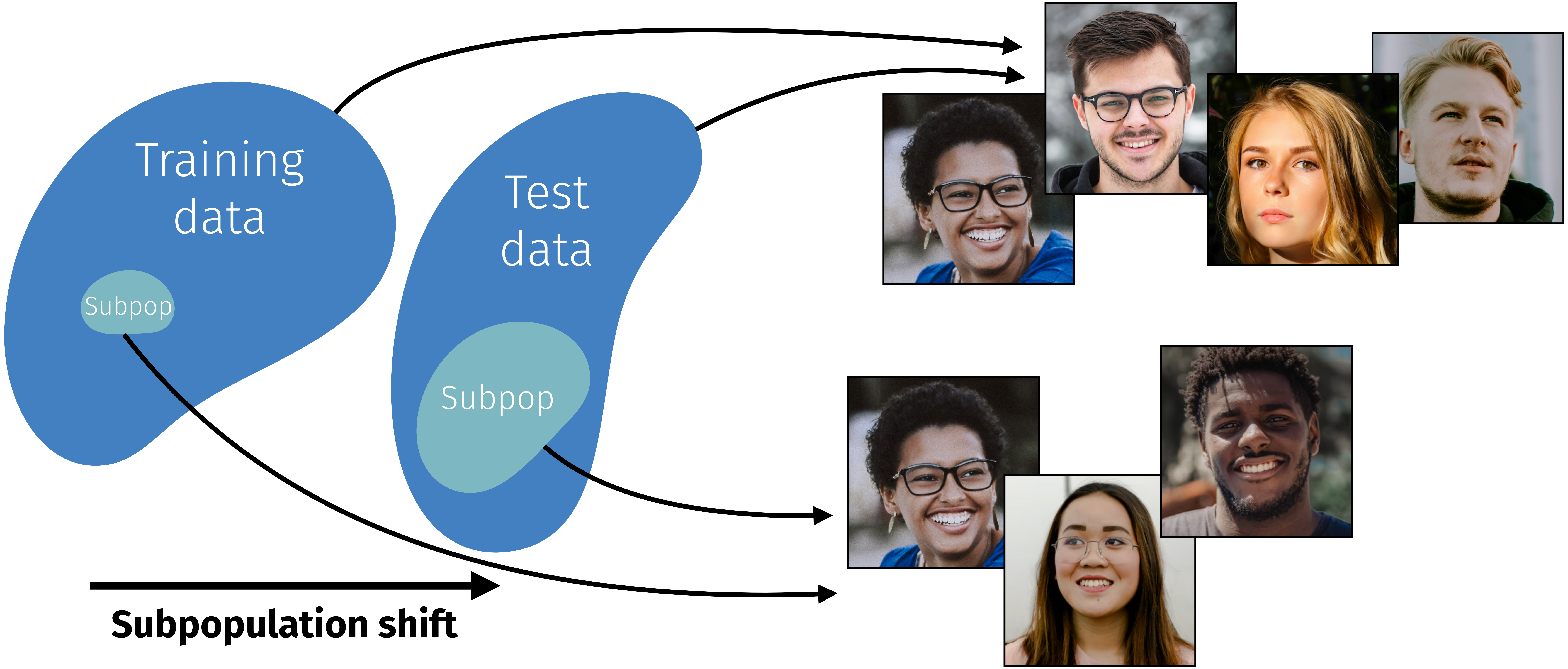
Fair machine learning

We want models that are robust across groups.



Fair machine learning

We want models that are robust across groups.

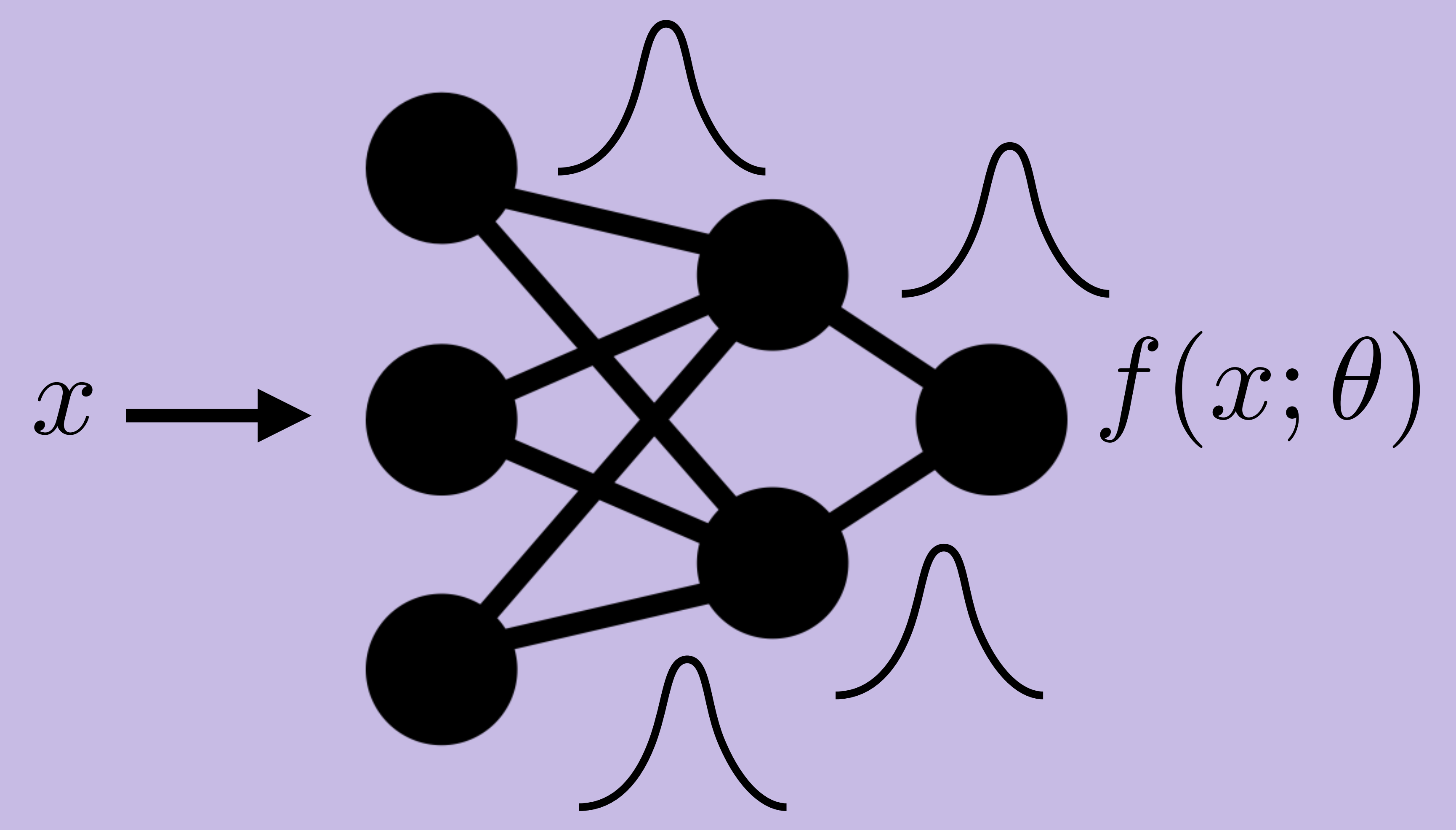


How can we train models that exhibit high group robustness?

A probabilistic approach

Find a **prior distribution** over **neural network parameters** that places high probability density on parameter values that **induce group robust classifiers**.

Probabilistic neural network

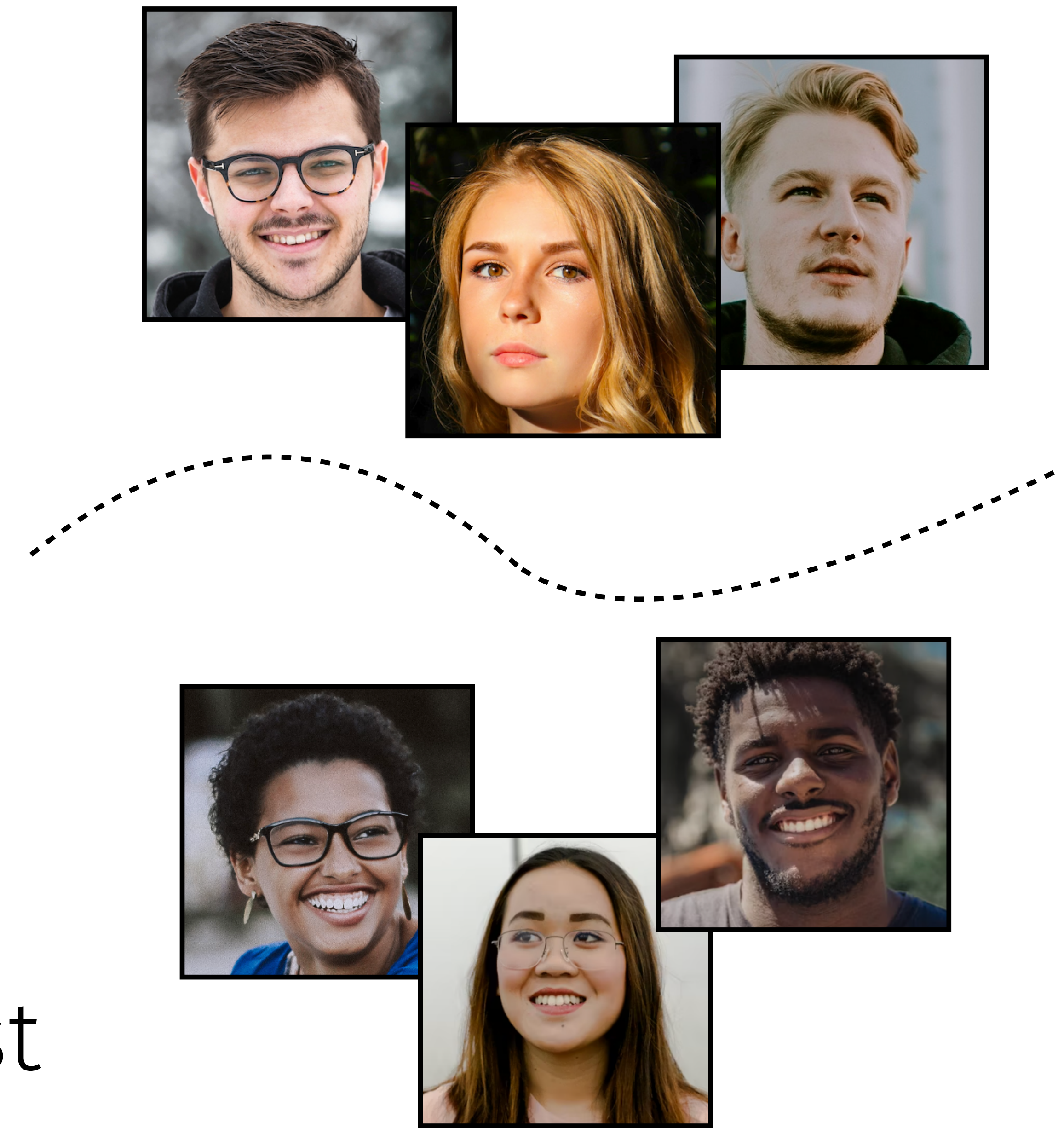


Group-aware priors (GAPs)

Goal: High accuracy across groups.

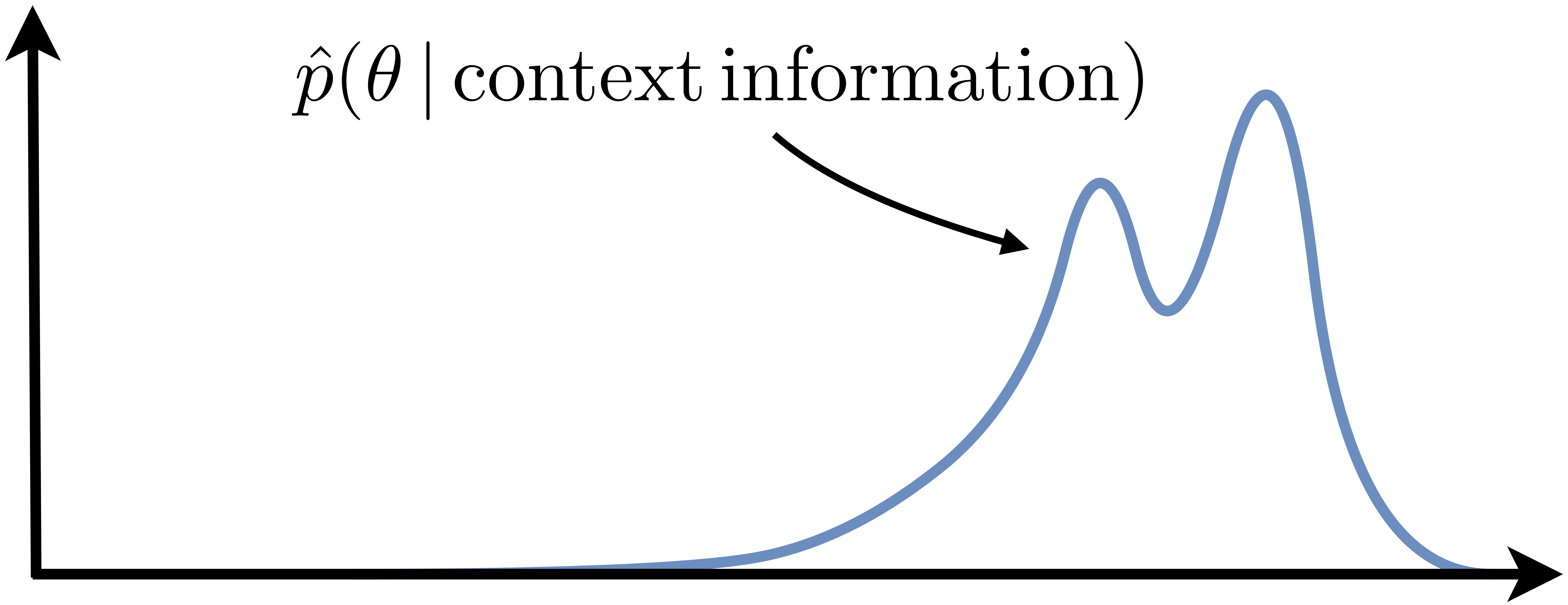
models that **are not** group-robust

models that **are** group-robust

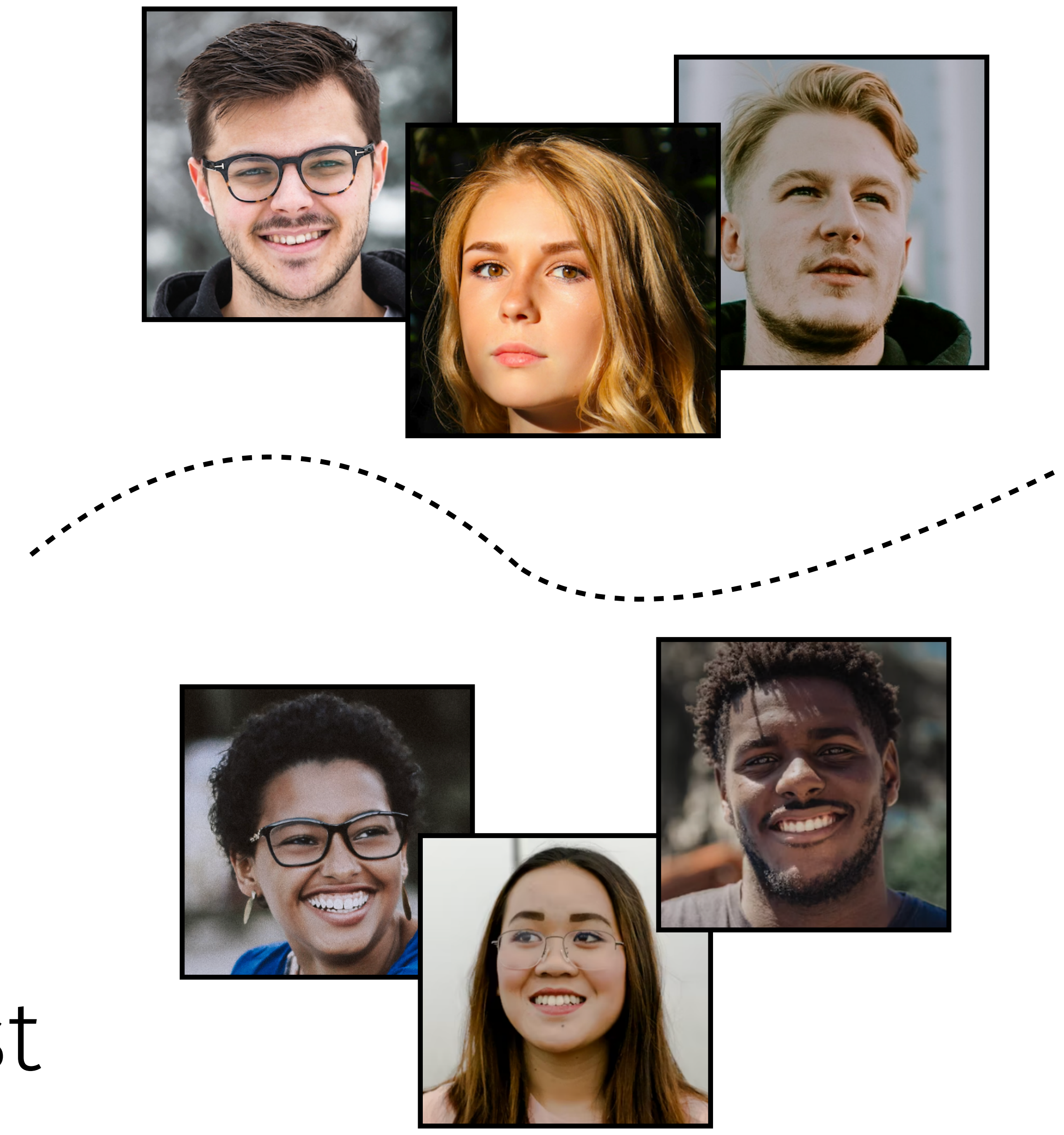


Group-aware priors (GAPs)

Goal: High accuracy across groups.



models that **are not** group-robust   models that **are** group-robust



How can we construct
data-driven group-aware priors?

Group-aware priors (GAPs)

Data-driven prior: $\hat{p}(\theta \,|\, \text{context information})$

Group-aware priors (GAPs)

Data-driven prior: $\hat{p}(\theta \,|\, \text{context information})$

Data-driven group-aware prior distribution

$$\hat{p}(\theta \mid \hat{z}; f, p_{\hat{X},\hat{Y}}) = \frac{\hat{p}(\hat{z} \,|\, \theta; f, p_{\hat{X},\hat{Y}})\, p(\theta)}{\hat{p}(\hat{z}; f, p_{\hat{X},\hat{Y}})}$$

Group-aware priors (GAPs)

Data-driven prior: $\hat{p}(\theta \mid \text{context information})$

Data-driven group-aware prior distribution

group-aware prior

$$\hat{p}(\theta \mid \hat{z}; f, p_{\hat{X},\hat{Y}}) = \frac{\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})\, p(\theta)}{\hat{p}(\hat{z}; f, p_{\hat{X},\hat{Y}})}$$

Group-aware priors (GAPs)

Data-driven prior: $\hat{p}(\theta \,|\, \text{context information})$

Data-driven group-aware prior distribution

auxiliary likelihood

group-aware prior

$$\hat{p}(\theta \mid \hat{z}; f, p_{\hat{X},\hat{Y}}) = \frac{\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})\, p(\theta)}{\hat{p}(\hat{z}; f, p_{\hat{X},\hat{Y}})}$$

Group-aware priors (GAPs)

Data-driven prior: $\hat{p}(\theta \mid \text{context information})$

Data-driven group-aware prior distribution

group-aware prior; auxiliary likelihood; base prior

$$\underbrace{\hat{p}(\theta \mid \hat{z}; f, p_{\hat{X},\hat{Y}})}_{\text{group-aware prior}} = \frac{\overbrace{\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})}^{\text{auxiliary likelihood}} \overbrace{p(\theta)}^{\text{base prior}}}{\hat{p}(\hat{z}; f, p_{\hat{X},\hat{Y}})}$$

Group-aware priors (GAPs)

Data-driven prior: $\hat{p}(\theta \,|\, \text{context information})$

Data-driven group-aware prior distribution

group-aware prior, auxiliary likelihood, base prior, marginal likelihood:

$$\underbrace{\hat{p}(\theta \mid \hat{z}; f, p_{\hat{X},\hat{Y}})}_{\text{group-aware prior}} = \frac{\overbrace{\hat{p}(\hat{z} \,|\, \theta; f, p_{\hat{X},\hat{Y}})}^{\text{auxiliary likelihood}} \, \overbrace{p(\theta)}^{\text{base prior}}}{\underbrace{\hat{p}(\hat{z}; f, p_{\hat{X},\hat{Y}})}_{\text{marginal likelihood}}}$$

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Auxiliary RV: $\hat{Z}$ ('achieving group robustness')

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Auxiliary RV: $\hat{Z}$ ('achieving group robustness')

Bernoulli likelihood:

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Auxiliary RV: $\hat{Z}$ ('achieving group robustness')

Bernoulli likelihood:

$$\hat{p}(\hat{z} = 1 \mid \theta; f, p_{\hat{X},\hat{Y}}) = \exp(-\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}}[c(\hat{X}, \hat{Y}, f, \theta)])$$

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Auxiliary RV: $\hat{Z}$ ('achieving group robustness')

Bernoulli likelihood:

$$\hat{p}(\hat{z} = 1 \mid \theta; f, p_{\hat{X},\hat{Y}}) = \exp(-\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}}[c(\hat{X}, \hat{Y}, f, \theta)])$$

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Auxiliary RV: $\hat{Z}$ ('achieving group robustness')

Bernoulli likelihood:

Cost function

$$\hat{p}(\hat{z} = 1 \mid \theta; f, p_{\hat{X},\hat{Y}}) = \exp(-\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}}[c(\hat{X}, \hat{Y}, f, \theta)])$$

Group-aware priors (GAPs)

Auxiliary likelihood: $\hat{p}(\hat{z} \mid \theta; f, p_{\hat{X},\hat{Y}})$

Auxiliary RV: $\hat{Z}$ ('achieving group robustness')

Bernoulli likelihood:

Distribution over context sets

Cost function

$$\hat{p}(\hat{z} = 1 \mid \theta; f, p_{\hat{X},\hat{Y}}) = \exp(-\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}}[c(\hat{X}, \hat{Y}, f, \theta)])$$

Group-aware priors (GAPs)

Specifying $p_{\hat{X},\hat{Y}}$:

Group-aware priors (GAPs)

Specifying $p_{\hat{X},\hat{Y}}$:

1. Assume access to a (small) dataset with group labels

Group-aware priors (GAPs)

Specifying $p_{\hat{X},\hat{Y}}$:

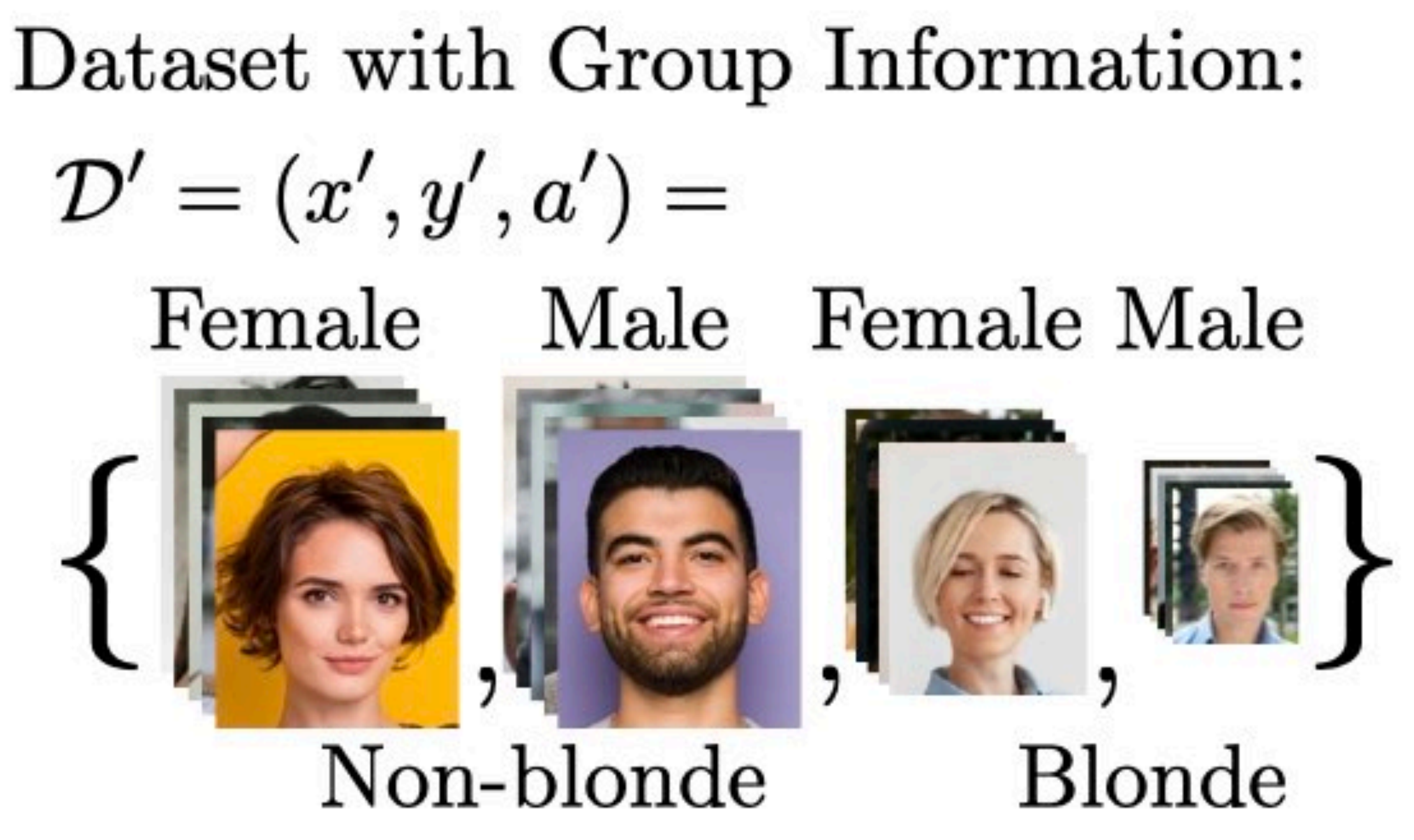
1. Assume access to a (small) dataset with group labels
2. Reweigh dataset by upsampling underrepresented groups

Group-aware priors (GAPs)

Specifying $p_{\hat{X},\hat{Y}}$:

1. Assume access to a (small) dataset with group labels

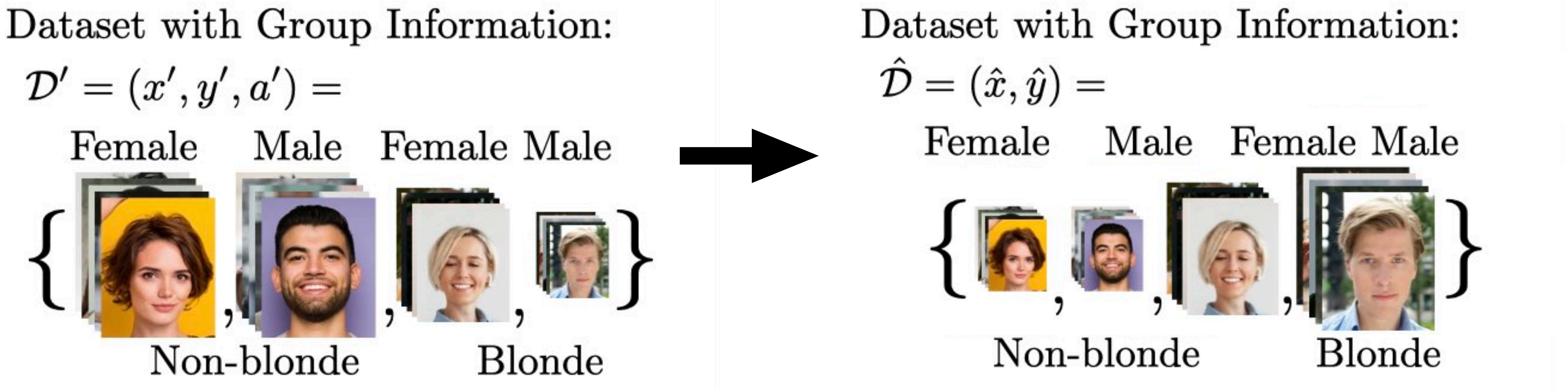
2. Reweigh dataset by upsampling underrepresented groups



Group-aware priors (GAPs)

Specifying $p_{\hat{X},\hat{Y}}$:

1. Assume access to a (small) dataset with group labels
2. Reweigh dataset by upsampling underrepresented groups

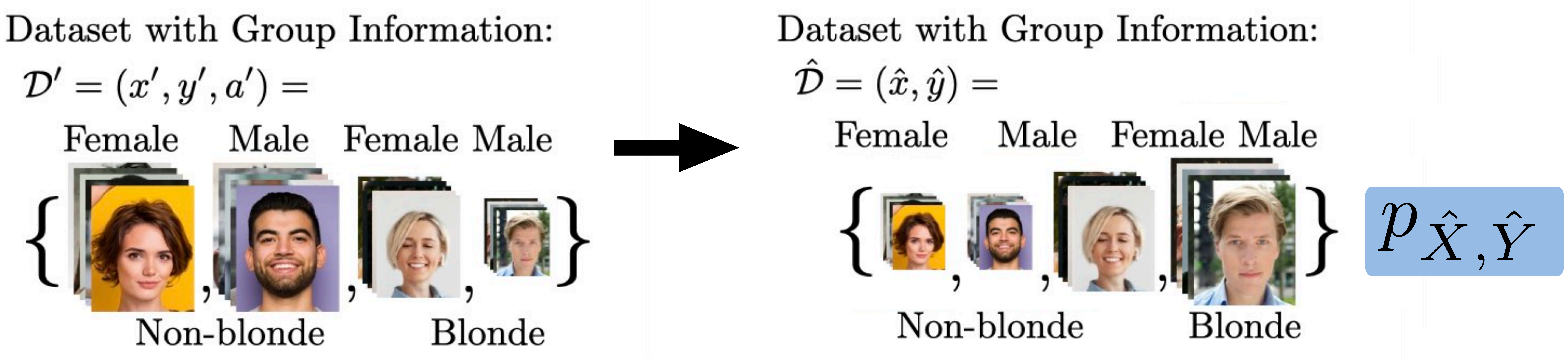


Group-aware priors (GAPs)

Specifying $p_{\hat{X},\hat{Y}}$:

1. Assume access to a (small) dataset with group labels

2. Reweigh dataset by upsampling underrepresented groups



Group-aware priors (GAPs)

Specifying $c(\hat{x}, \hat{y}, f, \theta)$:

Goal: Improve generalization for underrepresented group(s).

Group-aware priors (GAPs)

Specifying $c(\hat{x}, \hat{y}, f, \theta)$:

Goal: Improve generalization for underrepresented group(s).

1. Cross-entropy loss with parameter perturbation

$$c(\hat{x}, \hat{y}, f, \theta) \doteq \ell(\hat{y}, f(\hat{x}; \theta + \rho\epsilon(\theta)))$$

Group-aware priors (GAPs)

Specifying $c(\hat{x}, \hat{y}, f, \theta)$:

Goal: Improve generalization for underrepresented group(s).

1. Cross-entropy loss with parameter perturbation

$$c(\hat{x}, \hat{y}, f, \theta) \doteq \ell(\hat{y}, f(\hat{x}; \theta + \rho\epsilon(\theta)))$$

2. Worst-case perturbation

$$\epsilon(\theta, \hat{x}, \hat{y}) \doteq \perp \frac{\nabla_\theta \ell(\hat{y}, f(\hat{x}; \theta))}{\|\nabla_\theta \ell(\hat{y}, f(\hat{x}; \theta))\|_2}$$

Group-aware priors (GAPs)

Putting it all together:

Group-aware priors (GAPs)

Putting it all together:

- Let $\hat{z} = \{1, ..., 1\}$ (i.e., ‘group robustness achieved’)

Group-aware priors (GAPs)

Putting it all together:

- Let $\hat{z} = \{1, ..., 1\}$ (i.e., 'group robustness achieved')
- Auxiliary likelihood:

$$\hat{p}(\hat{z} = 1 \,|\, \theta; f, p_{\hat{X},\hat{Y}}) = \exp(-\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}}[\ell(\hat{y}, f(\hat{x}; \theta + \rho\epsilon(\theta)))])$$

Group-aware priors (GAPs)

Putting it all together:

- Let $\hat{z} = \{1, ..., 1\}$ (i.e., 'group robustness achieved')
- Auxiliary likelihood:

$$\hat{p}(\hat{z} = 1 \,|\, \theta; f, p_{\hat{X},\hat{Y}}) = \exp(-\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}}[\ell(\hat{y}, f(\hat{x}; \theta + \rho\epsilon(\theta)))])$$

Worst-case perturbation

Training with group-aware priors

Maximum a posteriori (MAP) estimation:

$$\max_{\theta} p(\theta \mid y_{\mathcal{D}}, x_{\mathcal{D}}, \hat{z}; f, p_{\hat{X},\hat{Y}})$$

Training with group-aware priors

Maximum a posteriori (MAP) estimation:

$$\max_{\theta} p(\theta \mid y_{\mathcal{D}}, x_{\mathcal{D}}, \hat{z}; f, p_{\hat{X},\hat{Y}})$$

↓

$$\min_{\theta} \Big\{ \underbrace{\sum_{n=1}^{N} \ell(y_{\mathcal{D}}^{(n)}, f(x_{\mathcal{D}}^{(n)}; \theta)) + \frac{\tau_\theta}{2} \|\theta - \mu\|_2^2}_{\text{standard } L_2\text{-regularized loss}} + \underbrace{\lambda \mathbb{E}_{p_{\hat{X},\hat{Y}}} [\ell(\hat{Y}, f(\hat{X}; \theta + \rho\epsilon(\theta)))]}_{\text{group robustness regularization}} \Big\}$$

Training with group-aware priors

Maximum a posteriori (MAP) estimation:

$$\max_{\theta} p(\theta \mid y_{\mathcal{D}}, x_{\mathcal{D}}, \hat{z}; f, p_{\hat{X},\hat{Y}})$$

$$\Downarrow$$

$$\min_{\theta} \Big\{ \underbrace{\sum_{n=1}^{N} \ell(y_{\mathcal{D}}^{(n)}, f(x_{\mathcal{D}}^{(n)}; \theta)) + \frac{\tau_\theta}{2} \|\theta - \mu\|_2^2}_{\text{standard } L_2\text{-regularized loss}} + \underbrace{\frac{\lambda}{S} \sum_{s=1}^{S} \ell(\hat{y}^{(s)}, f(\hat{x}^{(s)}; \theta + \rho\epsilon(\theta)))}_{\text{group robustness regularization}} \Big\}$$

Training with group-aware priors

Maximum a posteriori (MAP) estimation:

$$\max_{\theta} p(\theta \mid y_{\mathcal{D}}, x_{\mathcal{D}}, \hat{z}; f, p_{\hat{X},\hat{Y}})$$

$$\min_{\theta} \Big\{ \underbrace{\sum_{n=1}^{N} \ell(y_{\mathcal{D}}^{(n)}, f(x_{\mathcal{D}}^{(n)}; \theta)) + \frac{\tau_{\theta}}{2} \|\theta - \mu\|_2^2}_{\text{standard } L_2\text{-regularized loss}} + \underbrace{\frac{\lambda}{S} \sum_{s=1}^{S} \ell(\hat{y}^{(s)}, f(\hat{x}^{(s)}; \theta + \rho\epsilon(\theta)))}_{\text{group robustness regularization}} \Big\}$$

→ Amenable to stochastic optimization.

Training with group-aware priors

Step 1: Finetune pre-trained model with ERM

Dataset without Group Information:

$\mathcal{D} = (x, y) =$

Non-blonde Blonde

*Size of picture corresponds to size of group

Training with group-aware priors

Step 1: Finetune pre-trained model with ERM

Dataset without Group Information:

$\mathcal{D} = (x, y) =$

Non-blonde Blonde

*Size of picture corresponds to size of group

(1) Select Pretrained Model $f(\cdot\,;\theta)$

Training with group-aware priors

Step 1: Finetune pre-trained model with ERM

Dataset without Group Information:

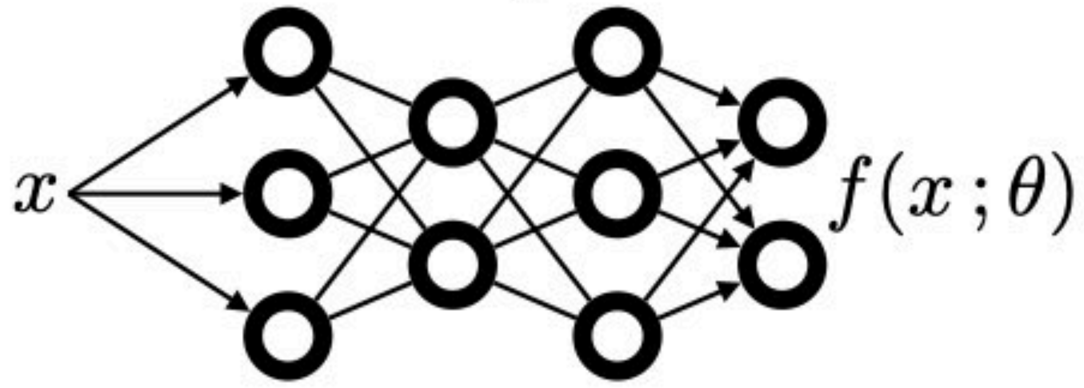
$\mathcal{D} = (x, y) =$

Non-blonde Blonde

*Size of picture corresponds to size of group

(1) Select Pretrained Model $f(\cdot\,; \theta)$

(2) Find: $\theta^* \doteq \arg\min_{\theta} \ell^{\mathrm{CE}}(y, f(x\,; \theta))$



Training with group-aware priors

Step 1: Finetune pre-trained model with ERM

Dataset without Group Information:

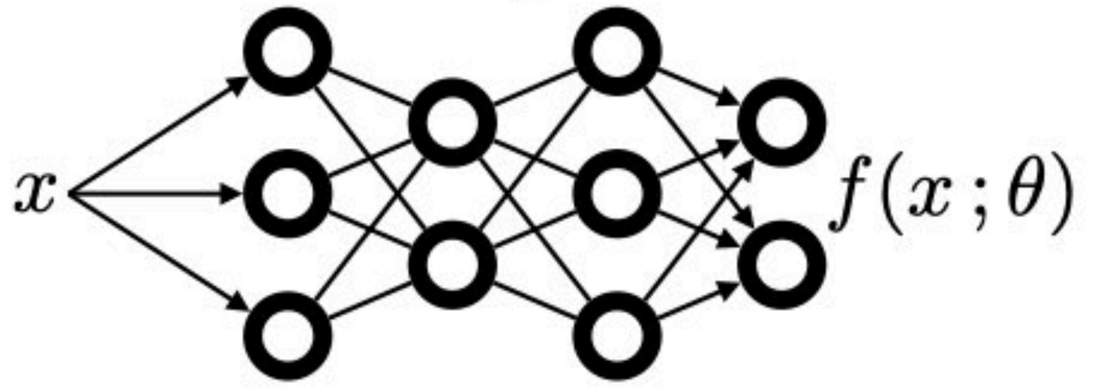
$\mathcal{D} = (x, y) =$

Non-blonde Blonde

*Size of picture corresponds to size of group

(1) Select Pretrained Model $f(\cdot\,;\theta)$

(2) Find: $\theta^* \doteq \arg\min_{\theta} \ell^{\text{CE}}(y, f(x\,;\theta))$

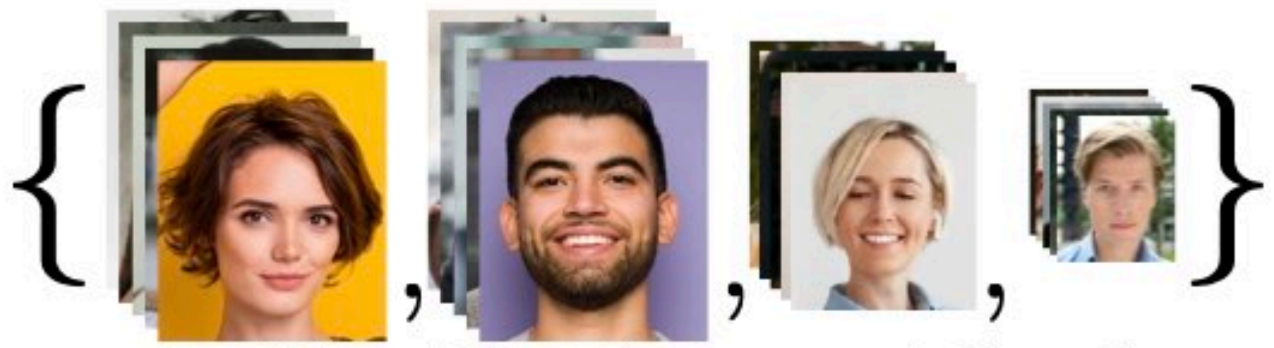


Step 2: Construct the group robustness prior

Dataset with Group Information:

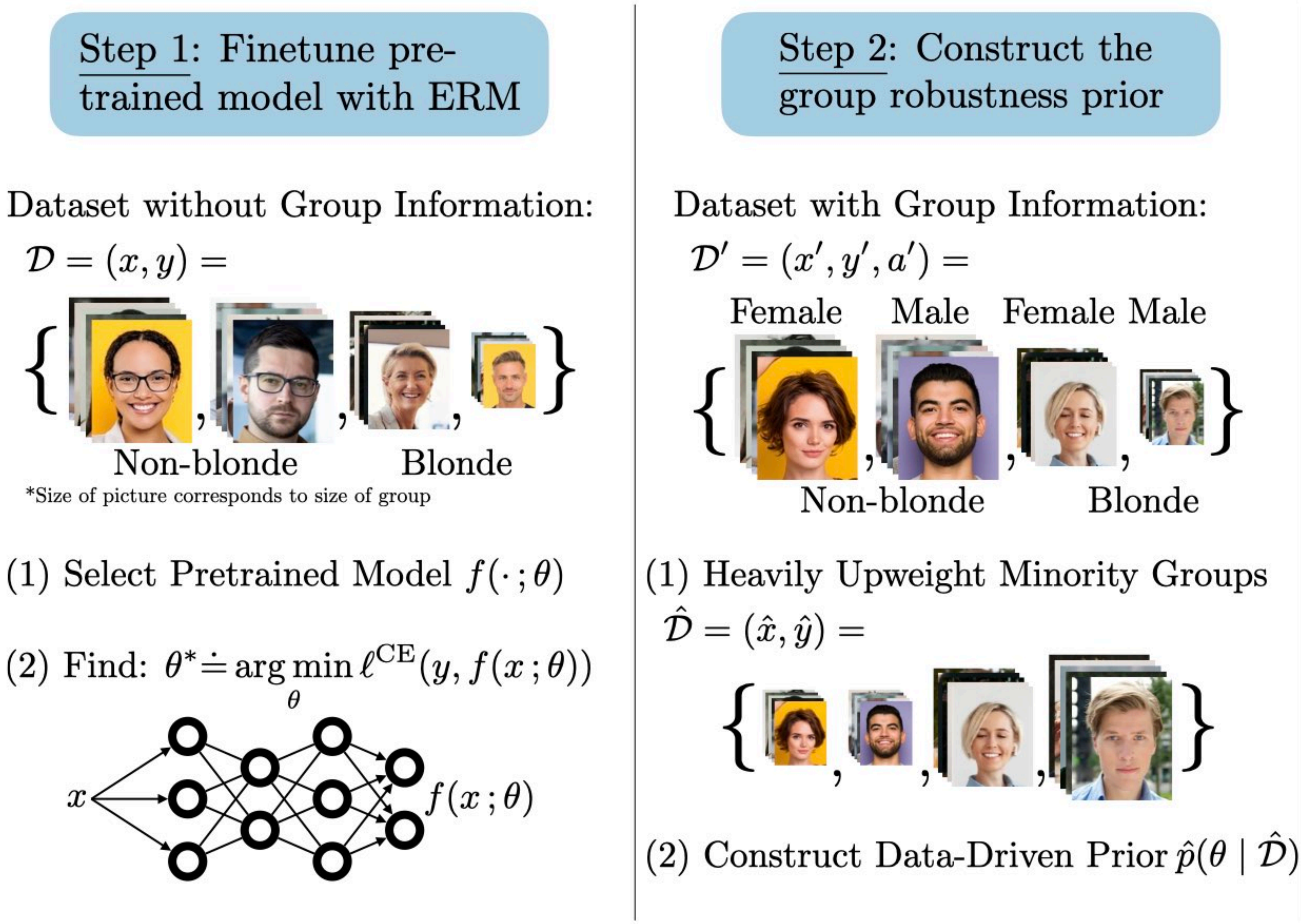
$\mathcal{D}' = (x', y', a') =$

Female Male Female Male

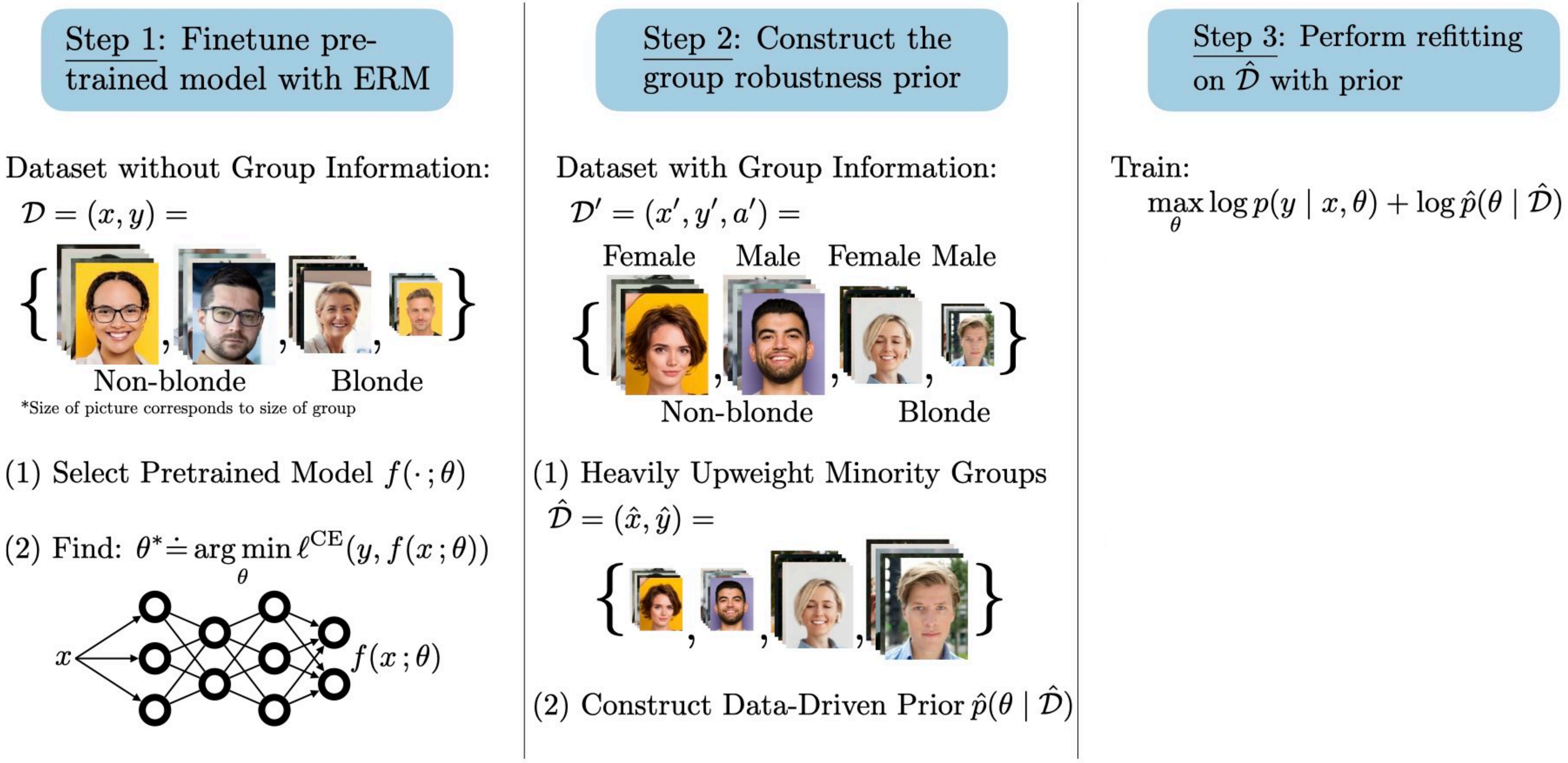


Non-blonde Blonde

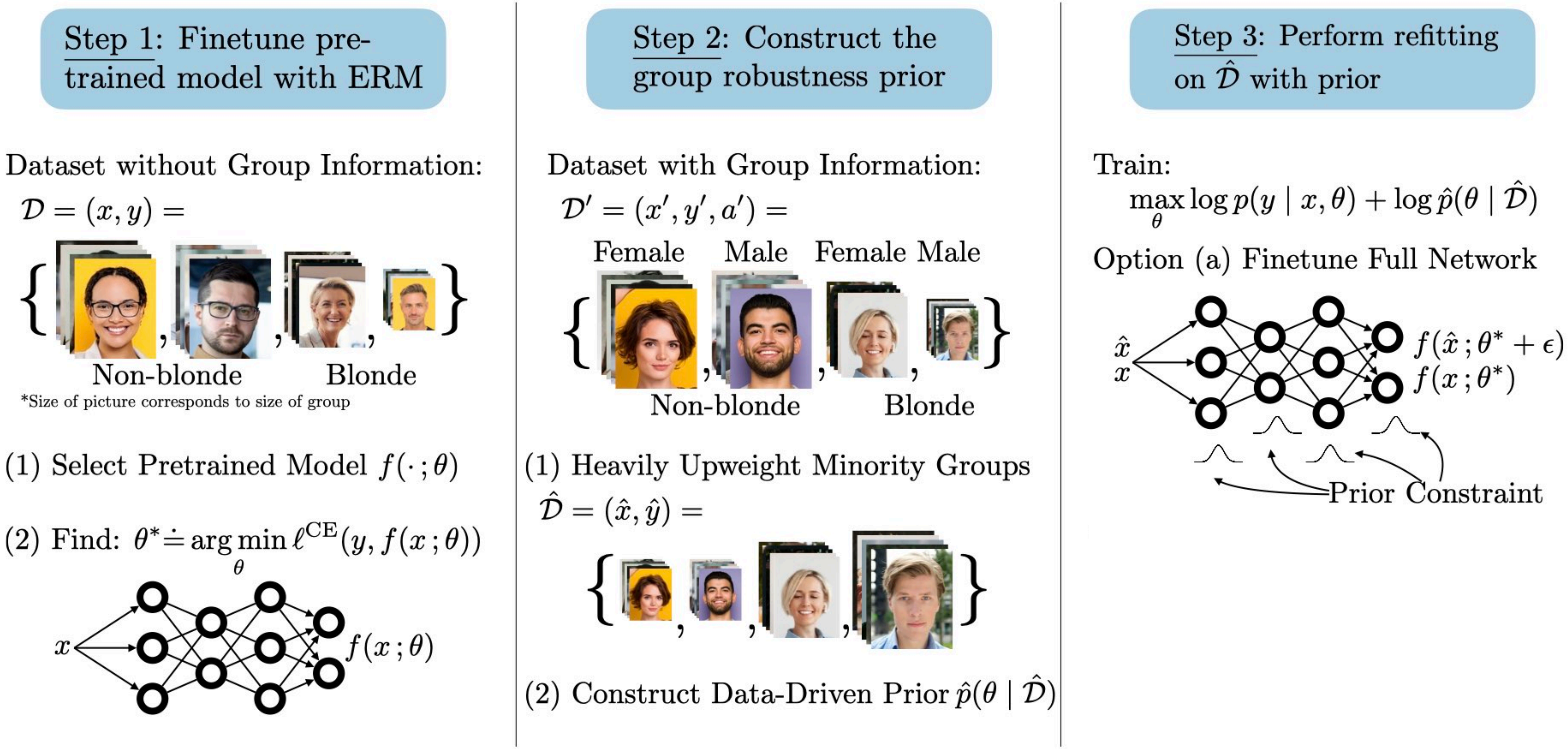
Training with group-aware priors



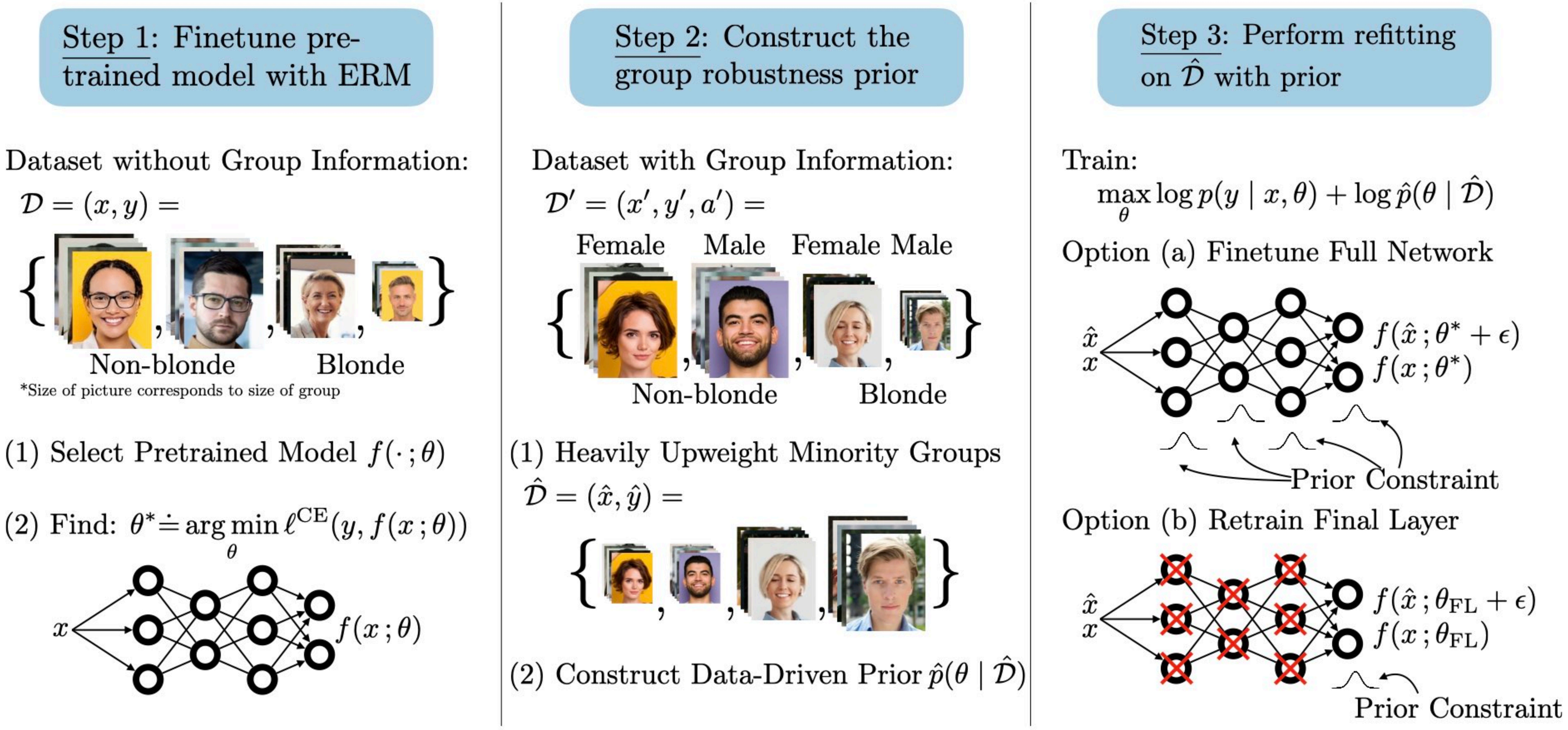
Training with group-aware priors



Training with group-aware priors



Training with group-aware priors



Evaluation: Worst-group accuracy

Waterbirds	
CelebA	
MultiNLI	

Evaluation: Worst-group accuracy

Contain **underrepresented groups**
and are exposed to **subpopulation shift**

Waterbirds	
CelebA	
MultiNLI	

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	ERM*
Waterbirds	74.9%
CelebA	49.9%
MultiNLI	65.9%

* ERM = Expected Risk Minimization (Vapnik, 1998)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	ERM*
Waterbirds	74.9% (98.1%)
CelebA	49.9% (95.3%)
MultiNLI	65.9% (82.8%)

* ERM = Expected Risk Minimization (Vapnik, 1998)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	ERM*		GAP
Waterbirds	74.9% (98.1%)	**+18.9%** →	93.8%
CelebA	49.9% (95.3%)	**+40.3%** →	90.2%
MultiNLI	65.9% (82.8%)	**+11.9%** →	77.8%

* ERM = Expected Risk Minimization (Vapnik, 1998)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	ERM*		GAP
Waterbirds	74.9% (98.1%)	+18.9% →	93.8% (95.6%)
CelebA	49.9% (95.3%)	+40.3% →	90.2% (91.5%)
MultiNLI	65.9% (82.8%)	+11.9% →	77.8% (82.5%)

* ERM = Expected Risk Minimization (Vapnik, 1998)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	G-DRO*
Waterbirds	
CelebA	
MultiNLI	

* DRO = Group Distributionally Robust Optimization (Sagawa et al., 2020)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	G-DRO*
Waterbirds	91.4% (93.5%)
CelebA	88.9% (92.2%)
MultiNLI	77.7% (81.9%)

* DRO = Group Distributionally Robust Optimization (Sagawa et al., 2020)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

	G-DRO*		GAP
Waterbirds	91.4% (93.5%)	+2.4% →	93.8% (95.6%)
CelebA	88.9% (92.2%)	+1.3% →	90.2% (91.5%)
MultiNLI	77.7% (81.9%)	+0.1% →	77.8% (82.5%)

* DRO = Group Distributionally Robust Optimization (Sagawa et al., 2020)

Evaluation: Worst-group accuracy

Contain **underrepresented groups** and are exposed to **subpopulation shift**

Only uses minimal group information!

	G-DRO*		GAP
Waterbirds	91.4% (93.5%)	+2.4% →	93.8% (95.6%)
CelebA	88.9% (92.2%)	+1.3% →	90.2% (91.5%)
MultiNLI	77.7% (81.9%)	+0.1% →	77.8% (82.5%)

* DRO = Group Distributionally Robust Optimization (Sagawa et al., 2020)

Evaluation: Full results

Table 1: **Average and Worst-Group Accuracy**.

Method	Group Info			Waterbirds		CelebA		MultiNLI	
	Tr.	Val.	Aux.	Worst	Average	Worst	Average	Worst	Average
ERM	N	N	N	74.9±1.0	98.1±0.0	46.9±1.3	95.3±0.0	65.9±0.1	82.8±0.0
JTT	N	Y	Y	86.7	93.3	81.1	88.0	72.6	78.6
CnC	N	Y	Y	88.5±0.2	90.9±0.1	88.8±0.5	89.9±0.3	—	—
SSA	N	Y	Y	89.0±0.3	92.2±0.5	89.8±0.8	**92.8±0.1**	76.6±0.4	79.9±0.5
DFR	N	Y	N	92.9±0.1	94.2±0.2	88.3±0.5	91.3±0.1	74.7±0.3	**82.1±0.1**
SUBG	Y	Y	N	89.1±0.5	—	85.6±1.0	—	68.9±0.4	—
G-DRO	Y	Y	N	91.4	93.5	88.9	**92.9**	**77.7**	81.4
GAP Last Layer	N	Y	N	**93.2±0.2**	**94.6±0.2**	**90.2±0.3**	91.7±0.2	74.3±0.2	81.9±0.0
GAP All Layers	N	Y	N	**93.8±0.1**	**95.6±0.1**	**90.2±0.3**	91.5±0.1	**77.8±0.6**	**82.5±0.1**

Evaluation: Full results

Table 1: **Average and Worst-Group Accuracy**.

Method	Group Info			Waterbirds		CelebA		MultiNLI	
	Tr.	Val.	Aux.	Worst	Average	Worst	Average	Worst	Average
ERM	N	N	N	74.9±1.0	98.1±0.0	46.9±1.3	95.3±0.0	65.9±0.1	82.8±0.0
JTT	N	Y	Y	86.7	93.3	81.1	88.0	72.6	78.6
CnC	N	Y	Y	88.5±0.2	90.9±0.1	88.8±0.5	89.9±0.3	—	—
SSA	N	Y	Y	89.0±0.3	92.2±0.5	89.8±0.8	**92.8±0.1**	76.6±0.4	79.9±0.5
DFR	N	Y	N	92.9±0.1	94.2±0.2	88.3±0.5	91.3±0.1	74.7±0.3	**82.1±0.1**
SUBG	Y	Y	N	89.1±0.5	—	85.6±1.0	—	68.9±0.4	—
G-DRO	Y	Y	N	91.4	93.5	88.9	**92.9**	**77.7**	81.4
GAP Last Layer	N	Y	N	**93.2±0.2**	**94.6±0.2**	**90.2±0.3**	91.7±0.2	74.3±0.2	81.9±0.0
GAP All Layers	N	Y	N	**93.8±0.1**	**95.6±0.1**	**90.2±0.3**	91.5±0.1	**77.8±0.6**	**82.5±0.1**

Evaluation: Full results

Table 1: **Average and Worst-Group Accuracy**.

Method	Group Info Tr.	Val.	Aux.	Waterbirds Worst	Waterbirds Average	CelebA Worst	CelebA Average	MultiNLI Worst	MultiNLI Average
ERM	N	N	N	74.9±1.0	98.1±0.0	46.9±1.3	95.3±0.0	65.9±0.1	82.8±0.0
JTT	N	Y	Y	86.7	93.3	81.1	88.0	72.6	78.6
CnC	N	Y	Y	88.5±0.2	90.9±0.1	88.8±0.5	89.9±0.3	—	—
SSA	N	Y	Y	89.0±0.3	92.2±0.5	89.8±0.8	**92.8±0.1**	76.6±0.4	79.9±0.5
DFR	N	Y	N	92.9±0.1	94.2±0.2	88.3±0.5	91.3±0.1	74.7±0.3	**82.1±0.1**
SUBG	Y	Y	N	89.1±0.5	—	85.6±1.0	—	68.9±0.4	—
G-DRO	Y	Y	N	91.4	93.5	88.9	**92.9**	**77.7**	81.4
GAP Last Layer	N	Y	N	**93.2±0.2**	**94.6±0.2**	**90.2±0.3**	91.7±0.2	74.3±0.2	81.9±0.0
GAP All Layers	N	Y	N	**93.8±0.1**	**95.6±0.1**	**90.2±0.3**	91.5±0.1	**77.8±0.6**	**82.5±0.1**

Evaluation: Full results

Table 1: **Average and Worst-Group Accuracy.**

Method	Group Info Tr.	Val.	Aux.	Waterbirds Worst	Waterbirds Average	CelebA Worst	CelebA Average	MultiNLI Worst	MultiNLI Average
ERM	N	N	N	74.9±1.0	98.1±0.0	46.9±1.3	95.3±0.0	65.9±0.1	82.8±0.0
JTT	N	Y	Y	86.7	93.3	81.1	88.0	72.6	78.6
CnC	N	Y	Y	88.5±0.2	90.9±0.1	88.8±0.5	89.9±0.3	—	—
SSA	N	Y	Y	89.0±0.3	92.2±0.5	89.8±0.8	**92.8±0.1**	76.6±0.4	79.9±0.5
DFR	N	Y	N	92.9±0.1	94.2±0.2	88.3±0.5	91.3±0.1	74.7±0.3	**82.1±0.1**
SUBG	Y	Y	N	89.1±0.5	—	85.6±1.0	—	68.9±0.4	—
G-DRO	Y	Y	N	91.4	93.5	88.9	**92.9**	**77.7**	81.4
GAP Last Layer	N	Y	N	**93.2±0.2**	**94.6±0.2**	**90.2±0.3**	91.7±0.2	74.3±0.2	81.9±0.0
GAP All Layers	N	Y	N	**93.8±0.1**	**95.6±0.1**	**90.2±0.3**	91.5±0.1	**77.8±0.6**	**82.5±0.1**

Summary

Summary

1. We designed a family of data-driven **group-aware priors** (GAPs) and constructed a simple prior instantiation.

Summary

1. We designed a family of data-driven **group-aware priors** (GAPs) and constructed a simple prior instantiation.

2. In practice, group-aware priors can be used as simple **add-on regularizers** for standard optimization objectives.

Summary

1. We designed a family of data-driven **group-aware priors** (GAPs) and constructed a simple prior instantiation.

2. In practice, group-aware priors can be used as simple **add-on regularizers** for standard optimization objectives.

3. Group-aware priors significantly **improve group robustness** under subpopulation shifts.

Thank you!

Mind the GAP: Improving Robustness to Subpopulation Shifts with Group-Aware Priors

Tim G. J. Rudner
@timrudner

Ya Shi Zhang
@andrew_yashi

Andrew Gordon Wilson
@andrewgwils

Julia Kempe
@KempeLab

NYU

Correspondence to
tim.rudner@nyu.edu

Paper:
timrudner.com/gap