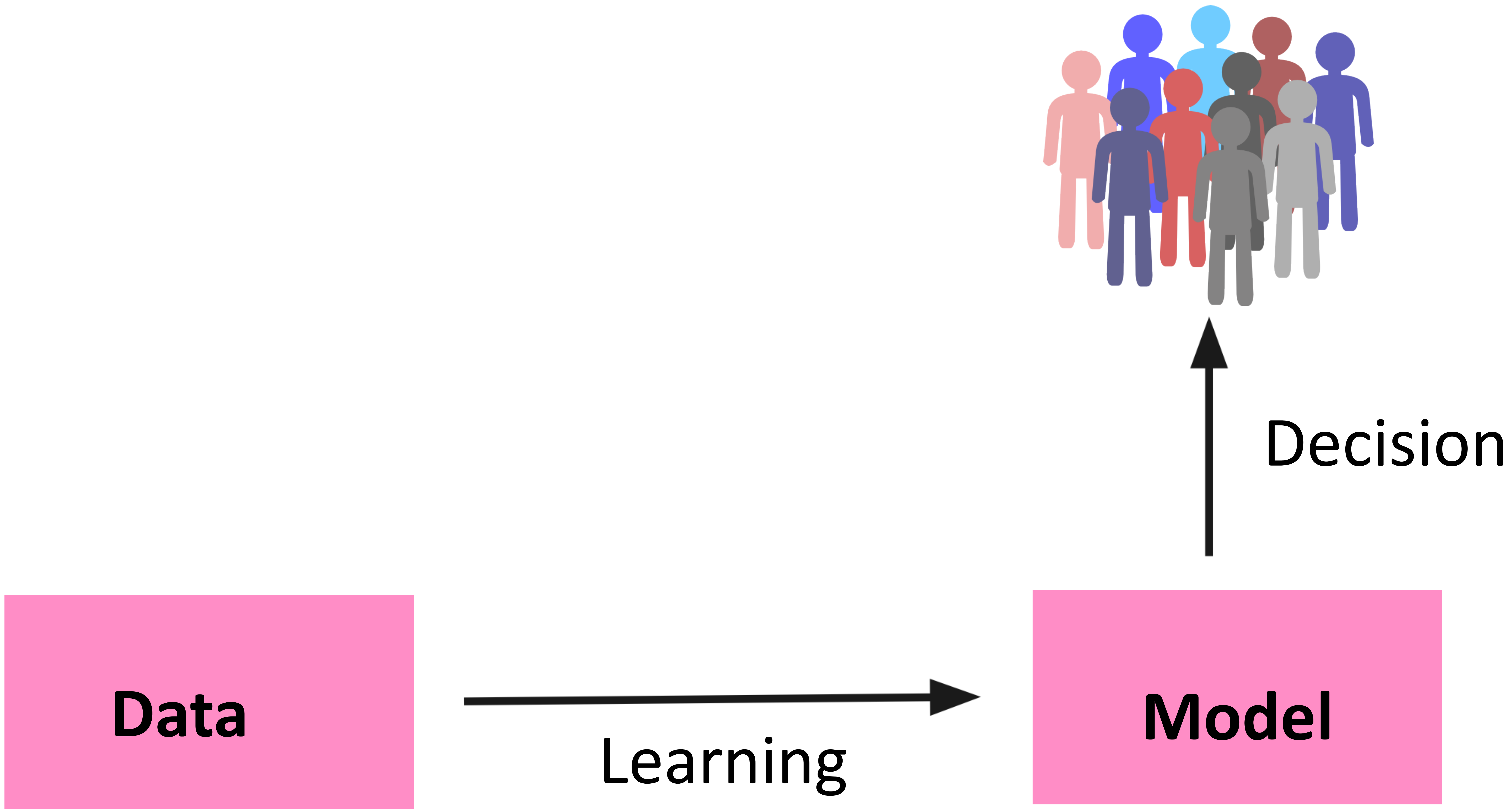
Regret Minimization with Performative Feedback

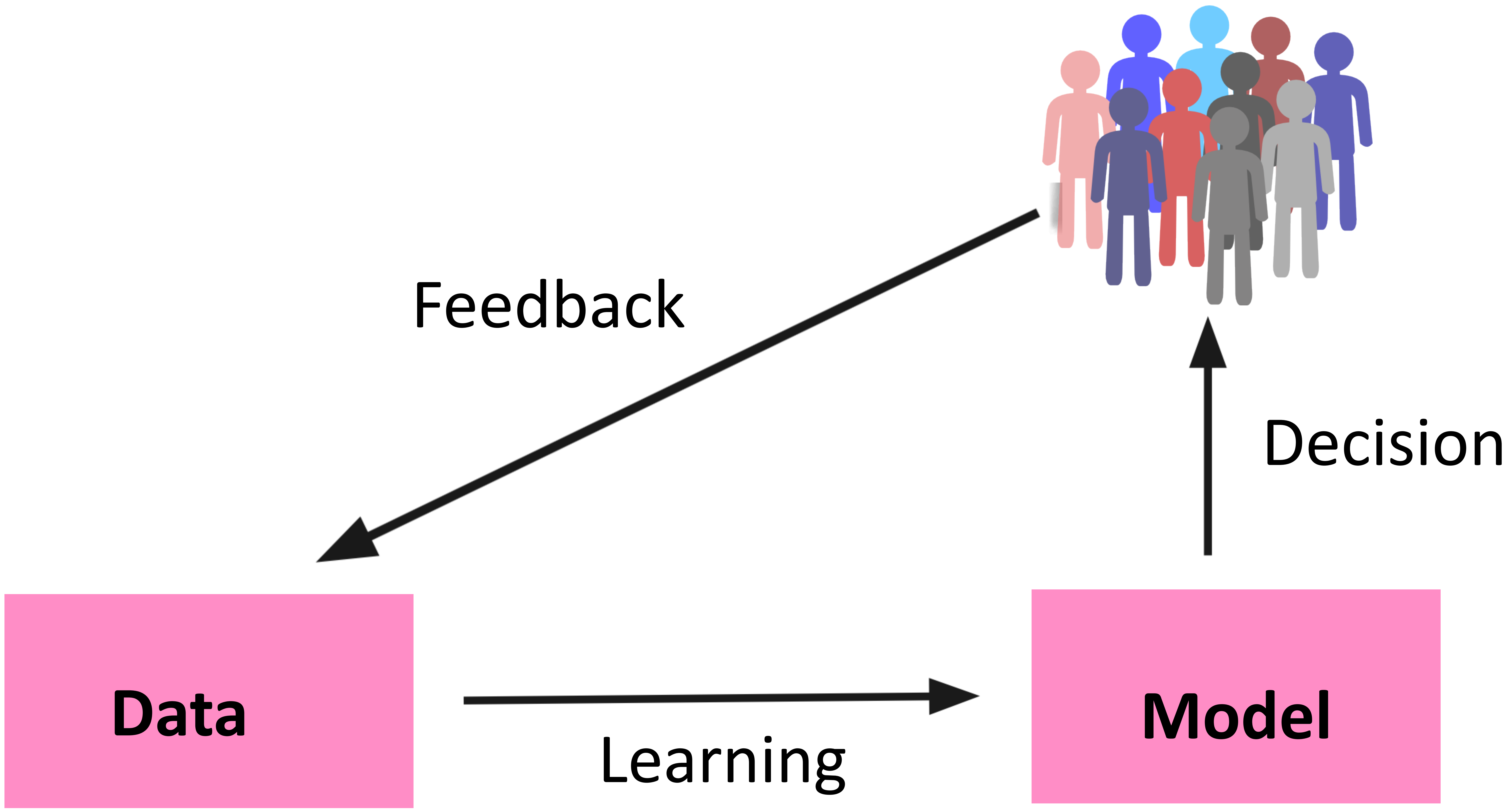
Meena Jagadeesan
University of California, Berkeley
joint work with
Tijana Zrnic
University of California, Berkeley
Celestine Mendler-Dünner
Max Planck Institute for Intelligent Systems

ICML 2022

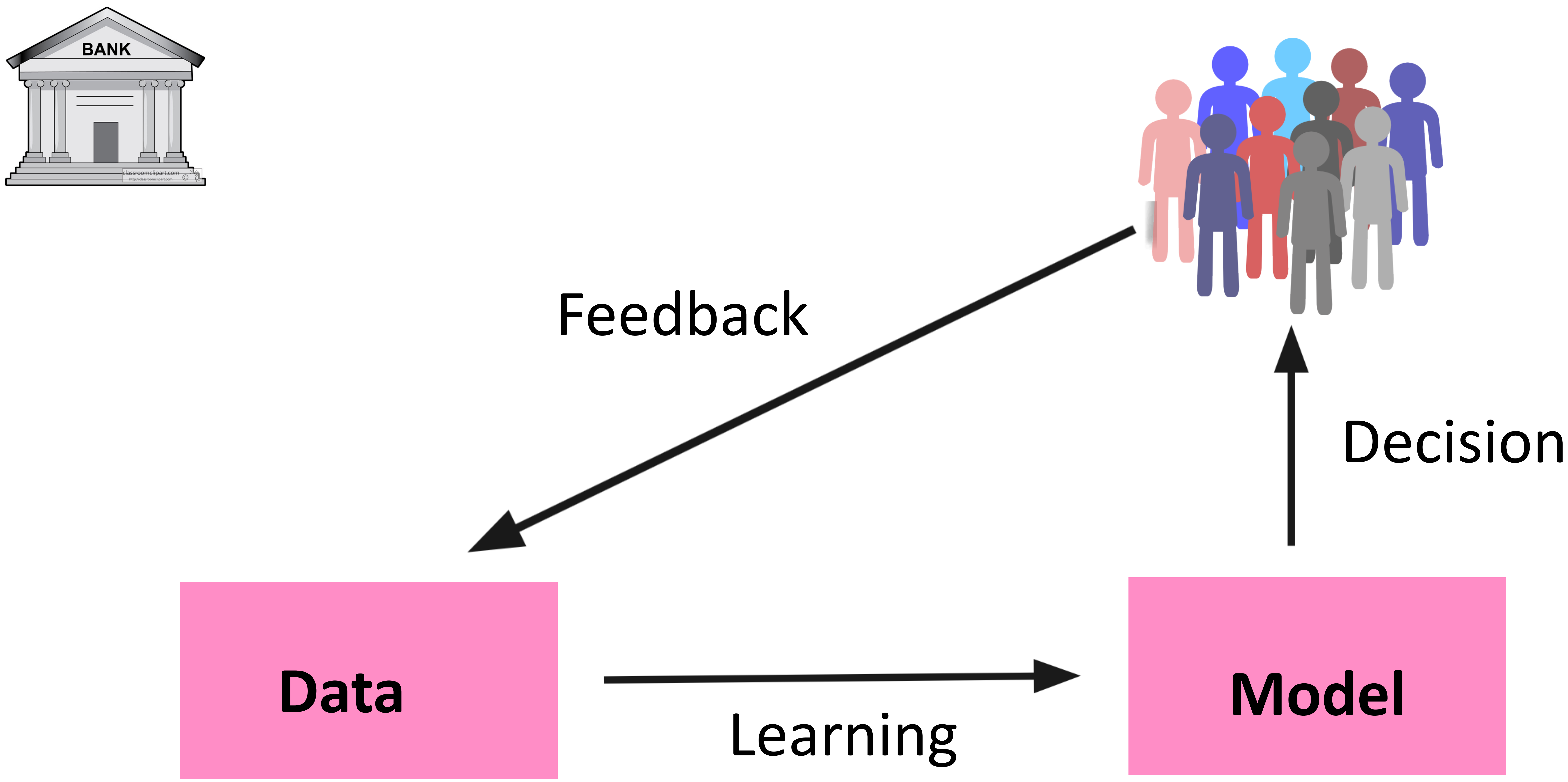
Static view of machine learning



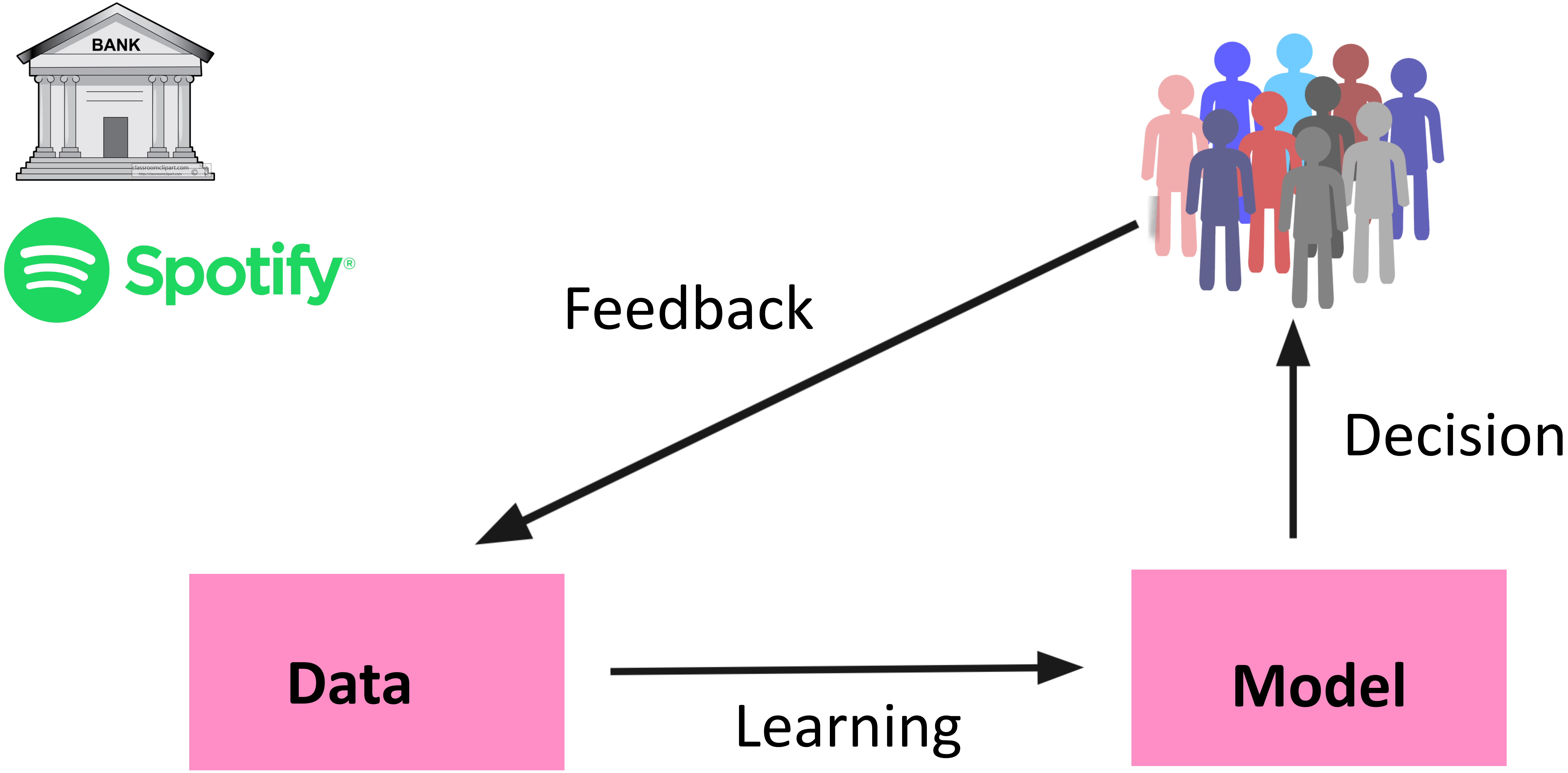
Machine learning with feedback effects



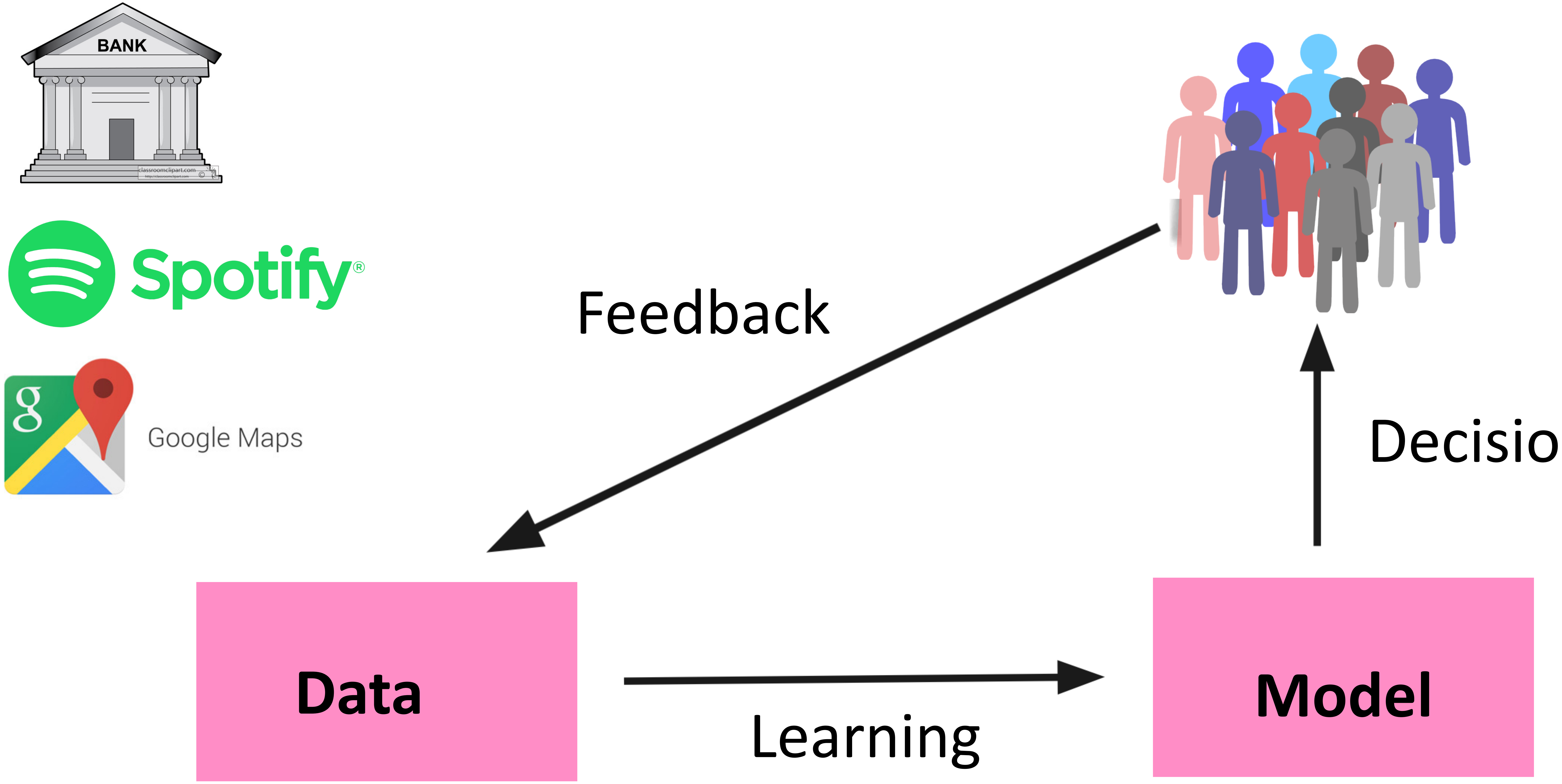
Machine learning with feedback effects



Machine learning with feedback effects



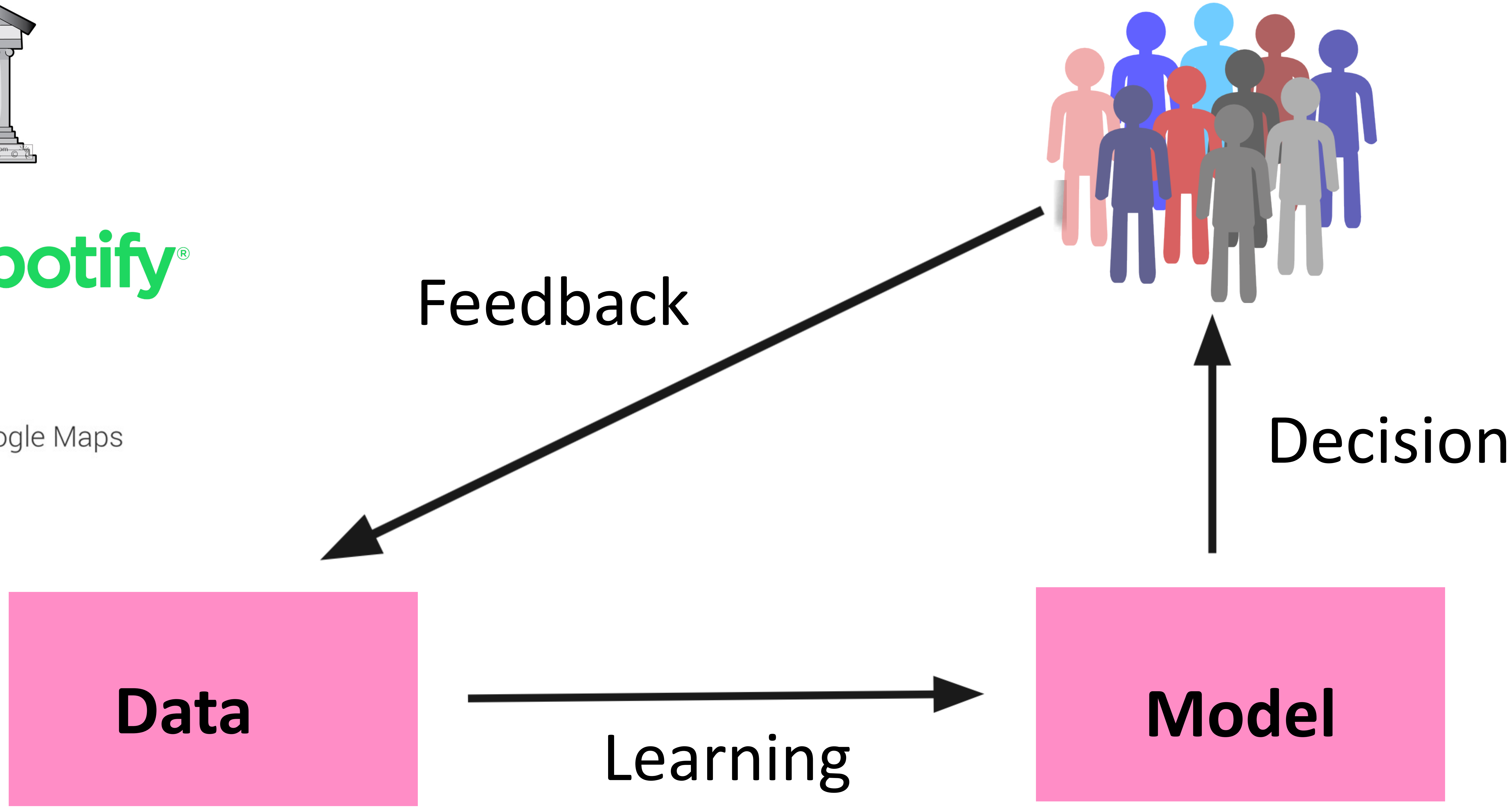
Machine learning with feedback effects



Machine learning with feedback effects

Spotify

Google Maps



Our contribution: Learning algorithms that perform well in the presence of feedback effects

(Perdomo, Zrnic, Mendler-Dünner, Hardt, 2020)

Performative prediction

Typical supervised learning: data $Z = (X, Y)$ distributed according to a fixed distribution $\mathcal{D}$

(Perdomo, Zrnic, Mendler-Dünner, Hardt, 2020)

Performative prediction

Typical supervised learning: data $Z = (X, Y)$ distributed according to a fixed distribution $\mathcal{D}$

Performative prediction: model induces a distribution shift in the data distribution.

- Each model f_θ induces possibly distinct distribution $Z = (X, Y)$ over observations.
- We call $\mathcal{D}(\theta)$ the distribution map.

(Perdomo, Zrnic, Mendler-Dünner, Hardt, 2020)

Performative prediction

Typical supervised learning: data $Z = (X, Y)$ distributed according to a fixed distribution $\mathcal{D}$

Performative prediction: model induces a distribution shift in the data distribution.

- Each model f_θ induces possibly distinct distribution $Z = (X, Y)$ over observations.
- We call $\mathcal{D}(\theta)$ the distribution map.

Objective: minimize *performative risk*:

$$\theta^* = \underset{\theta}{\operatorname{argmin}}\ \mathrm{PR}(\theta) \quad \mathrm{PR}(\theta) := \underset{Z \sim \mathcal{D}(\theta)}{\mathbb{E}} \ell(Z; \theta)$$

(Perdomo, Zrnic, Mendler-Dünner, Hardt, 2020)

Performative prediction

Typical supervised learning**:** data $Z = (X, Y)$ distributed according to a fixed distribution $\mathcal{D}$

Performative prediction: model induces a distribution shift in the data distribution.

- Each model f_θ induces possibly distinct distribution $Z = (X, Y)$ over observations.
- We call $\mathcal{D}(\theta)$ the distribution map.

Objective: minimize *performative risk*:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \, \mathrm{PR}(\theta) \quad \mathrm{PR}(\theta) := \underset{Z \sim \mathcal{D}(\theta)}{\mathbb{E}} \ell(Z; \theta)$$

The dependence on the model appears **twice**.

$\mathcal{D}(\theta)$ is **unknown**.

Our contributions

Approach: Gain information about distribution shifts through repeated model deployments

Our contributions

Approach: Gain information about distribution shifts through repeated model deployments

1. We establish a connection between performative prediction and bandits.

   Optimization in performative prediction $\approx$ bandit problem with **richer feedback**

2. Under smoothness assumptions, we design an algorithm whose regret scales with the complexity of the distribution map and *not* with the complexity of the performative risk.

3. We extend our results to linear distribution maps.

Performative optimization as online learning

- The learner needs to deploy different θ to explore the induced distributions $\mathcal{D}(\theta)$

- Natural to evaluate online sequence of deployments $\theta_1, \ldots, \theta_T$ via **performative regret**:

$$\text{Reg}(T) = \sum_{t=1}^{T} \Big(\mathbb{E}\text{PR}(\theta_t) - \text{PR}(\theta^*) \Big) \qquad \theta^* = \operatorname*{argmin}_{\theta} \text{PR}(\theta)$$

Performative optimization as online learning

- The learner needs to deploy different θ to explore the induced distributions $\mathcal{D}(\theta)$

- Natural to evaluate online sequence of deployments $\theta_1, \ldots, \theta_T$ via **performative regret**:

$$\mathrm{Reg}(T) = \sum_{t=1}^{T} \Big(\mathbb{E}\mathrm{PR}(\theta_t) - \mathrm{PR}(\theta^*) \Big) \quad \theta^* = \operatorname*{argmin}_{\theta} \mathrm{PR}(\theta)$$

"Baseline" bandits approach: Pull "arm" θ_t and observe reward $\widehat{\mathrm{PR}}(\theta_t)$.

Performative optimization as online learning

- The learner needs to deploy different θ to explore the induced distributions $\mathcal{D}(\theta)$

- Natural to evaluate online sequence of deployments $\theta_1, \dots, \theta_T$ via **performative regret**:

$$\mathrm{Reg}(T) = \sum_{t=1}^{T} \Big(\mathbb{E}\mathrm{PR}(\theta_t) - \mathrm{PR}(\theta^*) \Big) \qquad \theta^* = \operatorname*{argmin}_{\theta} \mathrm{PR}(\theta)$$

“Baseline” bandits approach: Pull “arm” θ_t and observe reward $\widehat{\mathrm{PR}}(\theta_t)$.

Main insight: performative settings exhibit **richer feedback** than bandit feedback.

- We observe **samples** from $\mathcal{D}(\theta_t)$, not just bandit feedback about performative risk.
- Can find θ^* with less exploration than bandit baselines.

Key insight: tighter confidence bounds

Conclusion and Future Work

- Deploying a model can induce a performative distribution shift on the population.
- Learner needs to deploy models online to find one $\approx$ with low induced risk
- Regret minimization with performative feedback $\approx$ bandit problem with richer feedback
- Performative feedback requires **less exploration** to find a good solution.

Future work: Leverage bandit tools for performative prediction more generally.

Thank you.