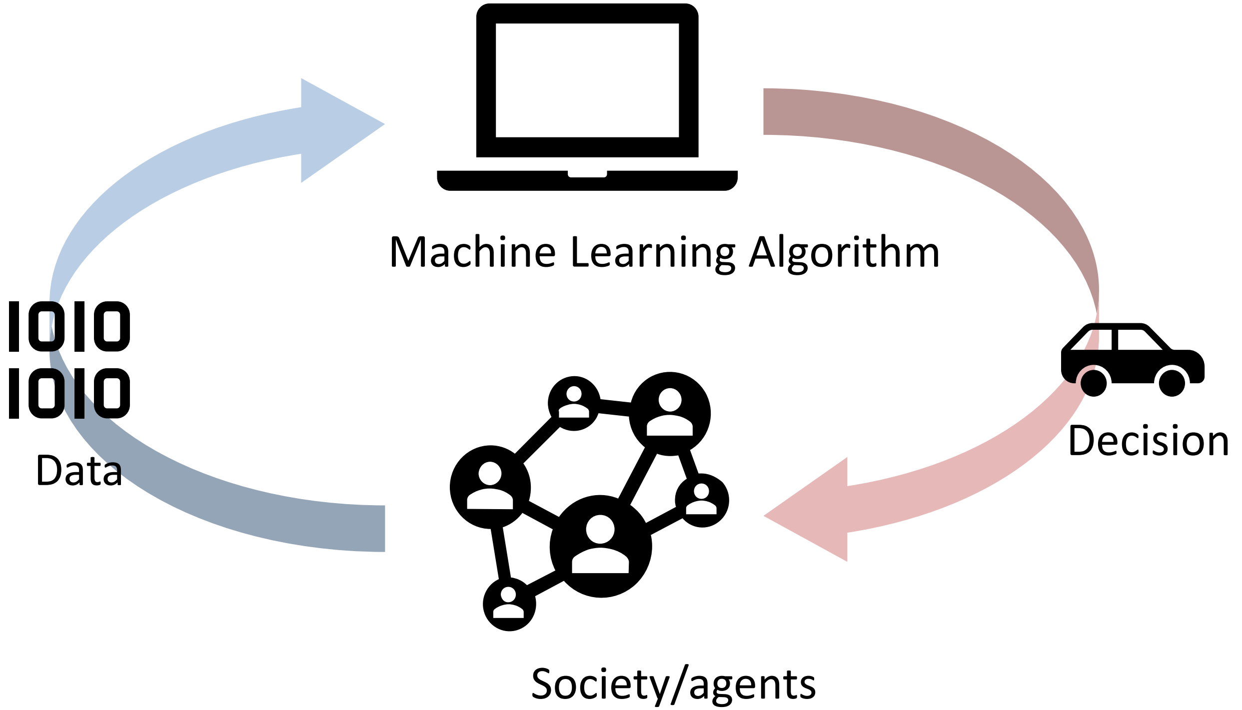
MULTI-AGENT PERFORMATIVE PREDICTION

Georgios Piliouras, Singapore University of Technology & Design

Fang-Yi Yu, George Mason University

Prediction is part of a broader system

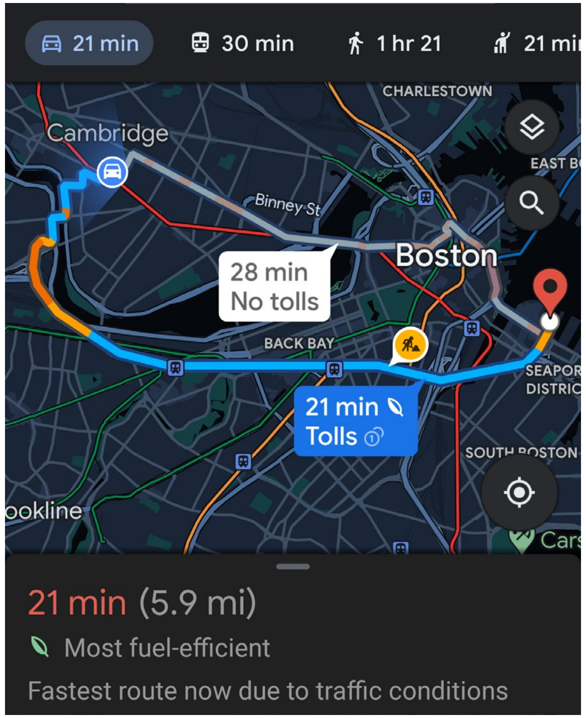
- When predictions support decisions, we fundamentally change the distribution of future data. [JZMH20]



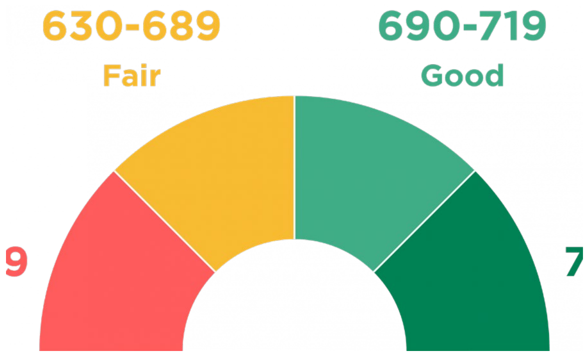
Perdomo, Juan, et al. "Performative prediction." ICML 2020.

Performative predictions are everywhere

Traffic prediction



Credit score



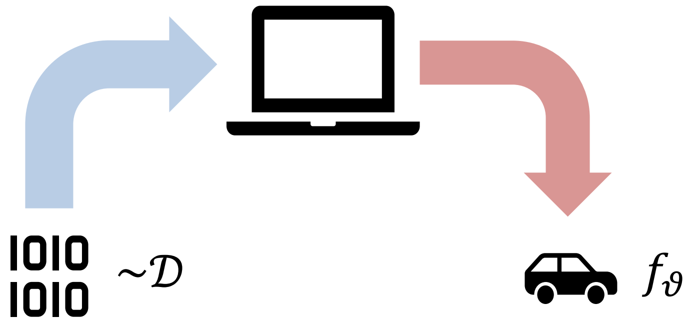
wildlife poaching predictions

Performative prediction

Supervised learning

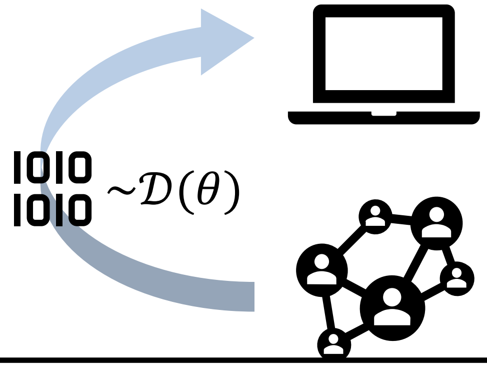
- Data $(x, y) \in \mathcal{X} \times \mathcal{Y}$ sampled from a fixed distribution $\mathcal{D}$
- Expected loss of a model f_ϑ with $\vartheta \in \Theta$

  $\ell(\vartheta; \mathcal{D}) \coloneqq \mathbb{E}_{(x,y)\sim\mathcal{D}}[\ell(f_\vartheta(x), y)]$



Performative prediction

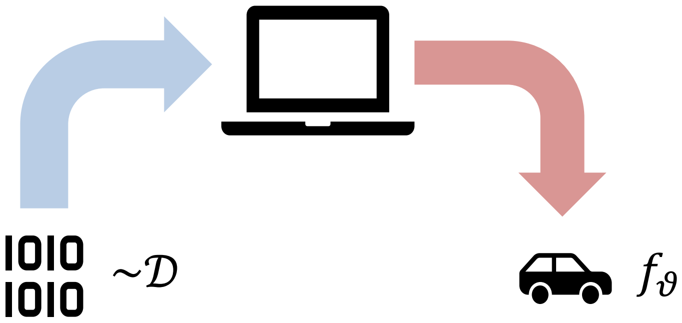
- Given $\theta \in \Theta$, data (x, y) is sampled from a distribution map $\mathcal{D}(\theta)$



Performative prediction

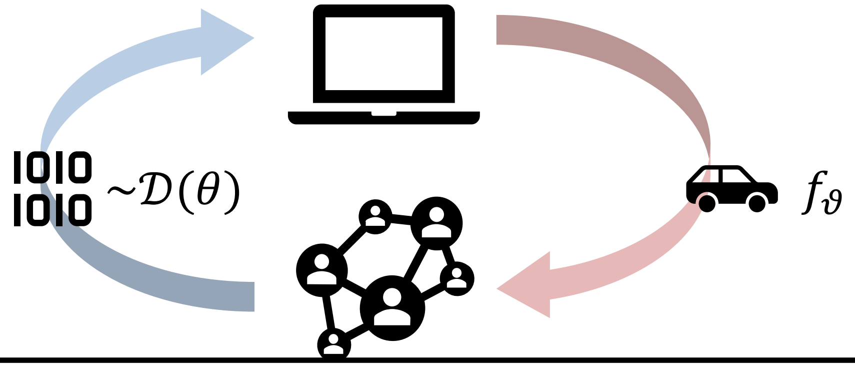
Supervised learning

- Data $(x, y) \in \mathcal{X} \times \mathcal{Y}$ sampled from a fixed distribution $\mathcal{D}$
- Expected loss of a model f_ϑ with $\vartheta \in \Theta$
  $\ell(\vartheta; \mathcal{D}) \coloneqq \mathbb{E}_{(x,y)\sim\mathcal{D}}[\ell(f_\vartheta(x), y)]$



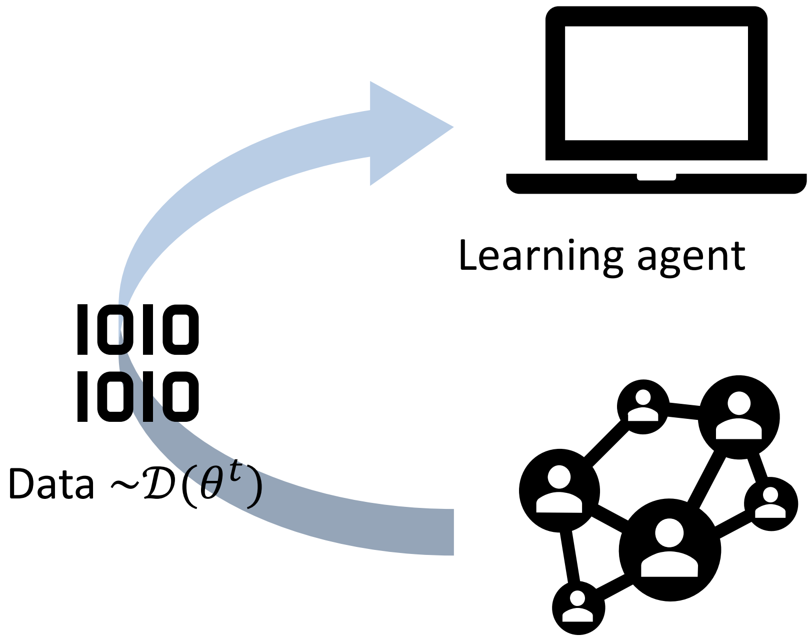
Performative prediction

- Given $\theta \in \Theta$, data (x, y) is sampled from a distribution map $\mathcal{D}(\theta)$
- Expected loss of a predictive model $\vartheta \in \Theta$ on deployed model $\theta \in \Theta$
  $\ell(\vartheta; \mathcal{D}(\theta)) := \mathbb{E}_{(x,y)\sim\mathcal{D}(\theta)}[\ell(f_\vartheta(x), y)]$



Dynamics of learning

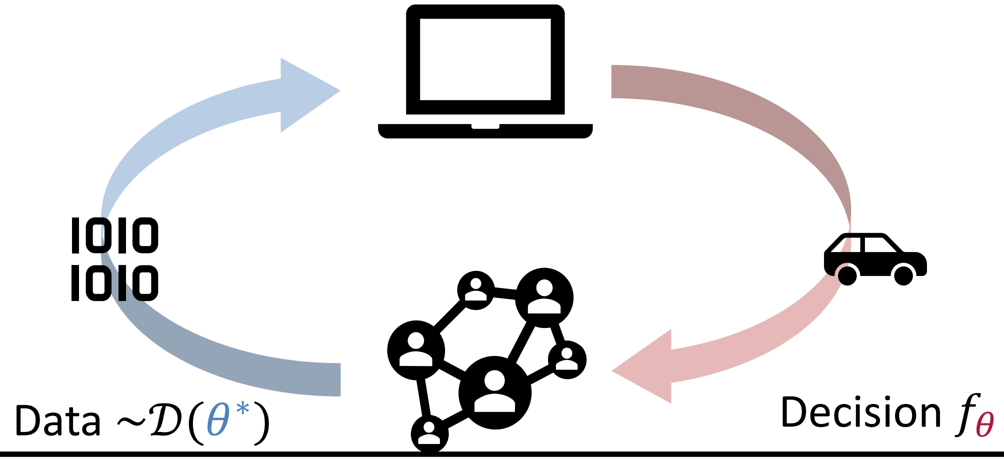
Since the distribution map $\mathcal{D}(\cdot)$ is unknown, the learning agent may iteratively access samples of $\mathcal{D}(\theta^t)$



Dynamics of learning

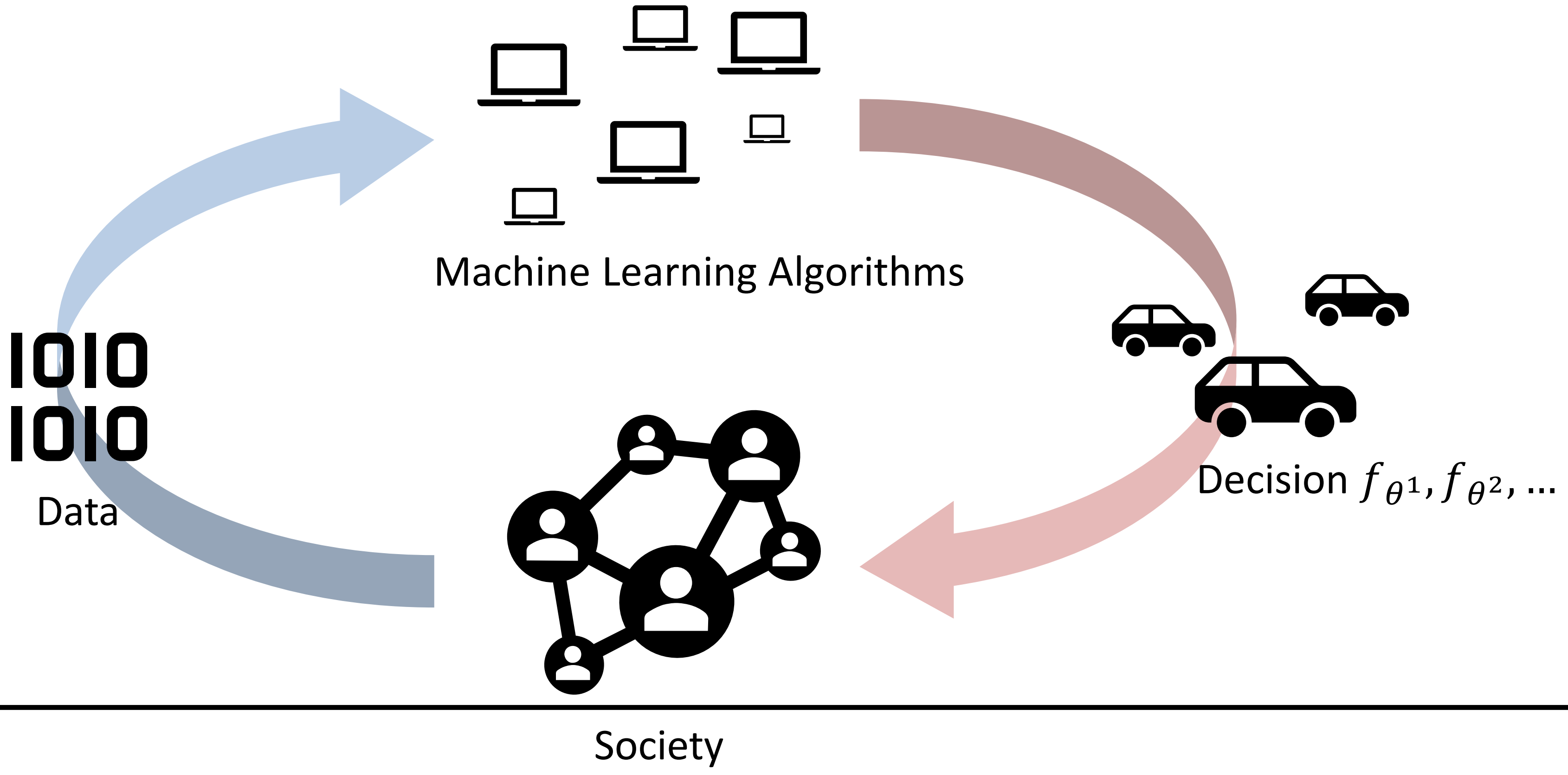
- The learning agent iteratively access samples of $\mathcal{D}(\theta^t)$ and improve the model to θ^{t+1}
- Fixed point: the model is optimal for distribution that it induces => performative stable point θ^*

$$\ell\big(\theta^*; \mathcal{D}(\theta^*)\big) = \min_{\theta} \ell\big(\theta; \mathcal{D}(\theta^*)\big)$$



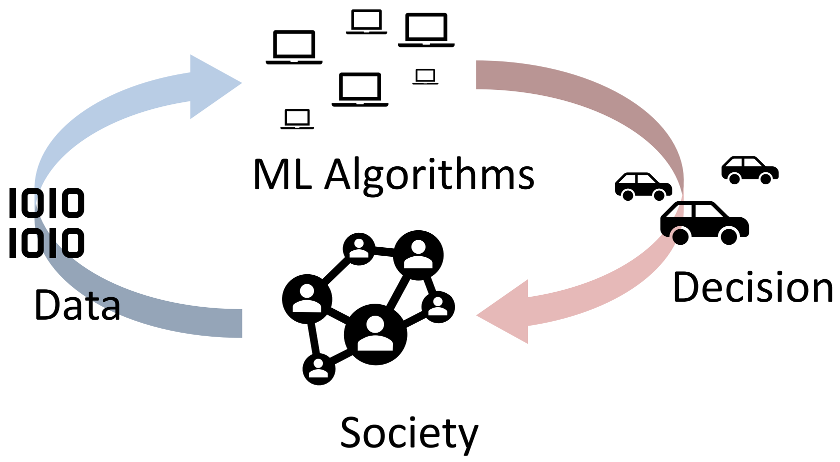
Multi-agent performative prediction

- Multiple predictive agents making decisions that collectively influence the distribution of future data.



Contributions

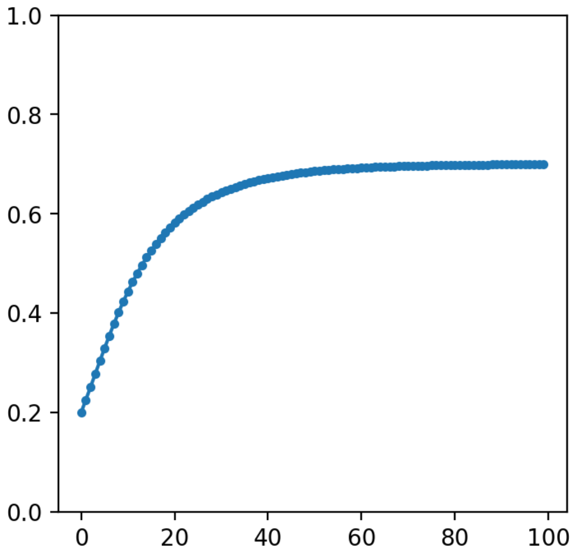
- Formalize the concept of multi agent performative predictions and a novel solution concept, multi agent performative stability
- Main result: a threshold result on RL algorithms
  - Convergence for small learning rate
  - Li-Yorke chaos when too much influence



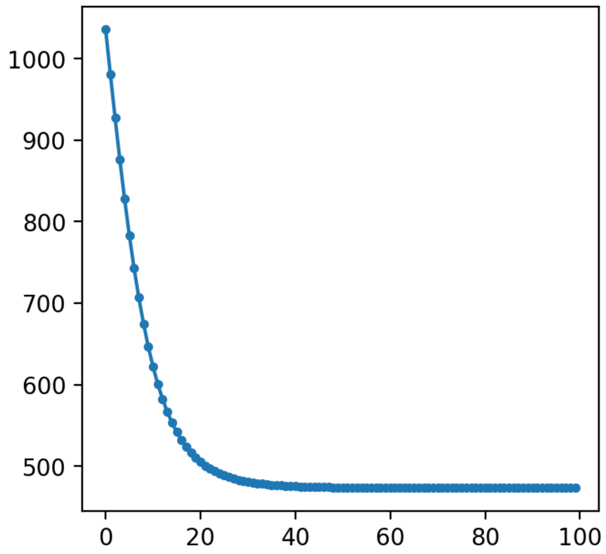
Georgios Piliouras, Fang-Yi Yu. “Multi-agent Performative Prediction: from Global Stability and Optimality to Chaos." under review 2022

Simulation

Small learning rate (η)

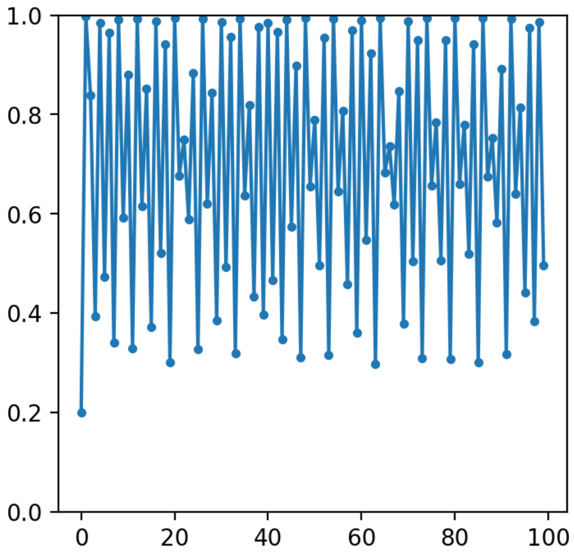


Model θ^t

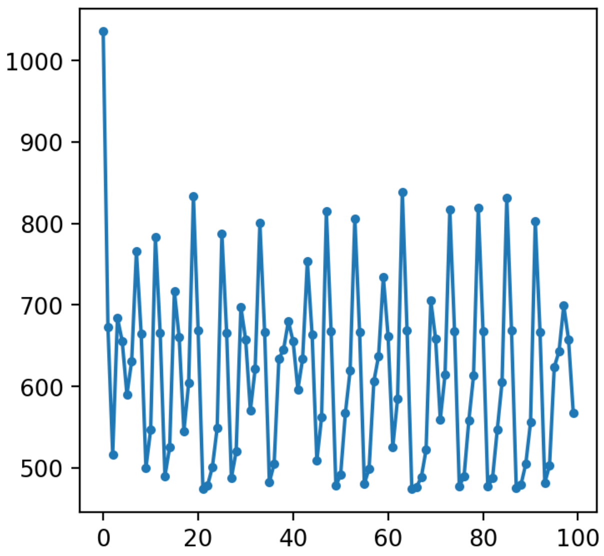


Loss $\ell\left(\theta^t; \mathcal{D}(\theta^t)\right)$

Large influence ($n\lambda$)



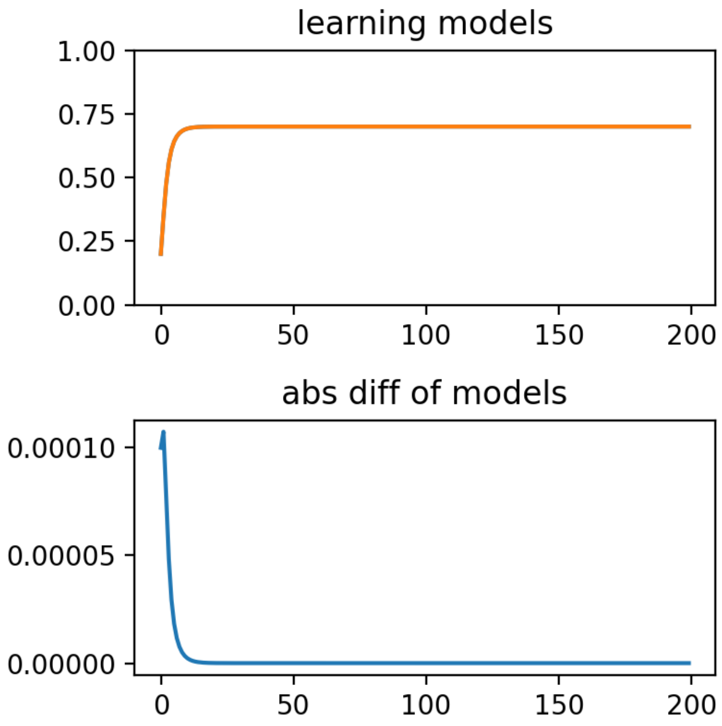
Model θ^t



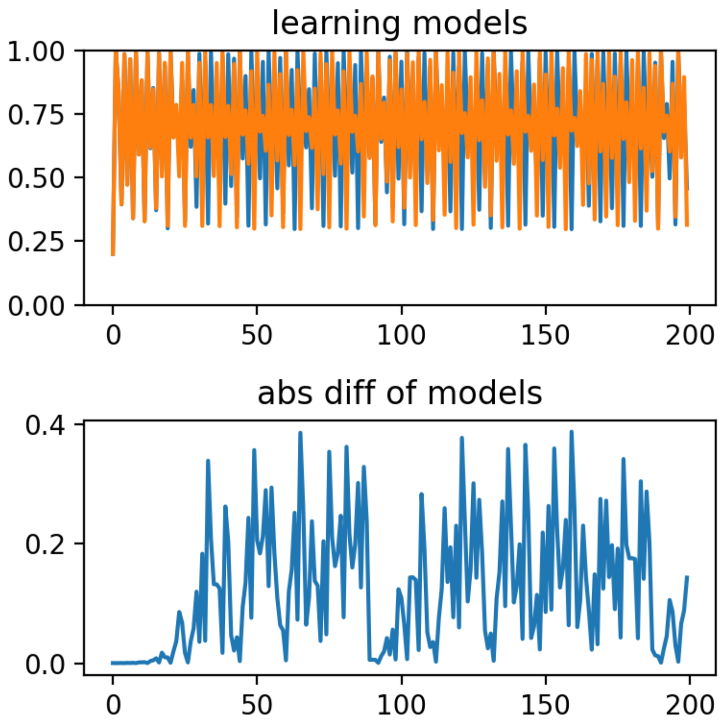
Loss $\ell\left(\theta^t; \mathcal{D}(\theta^t)\right)$

Simulation

Small learning rate (η)



Large influence ($n\lambda$)

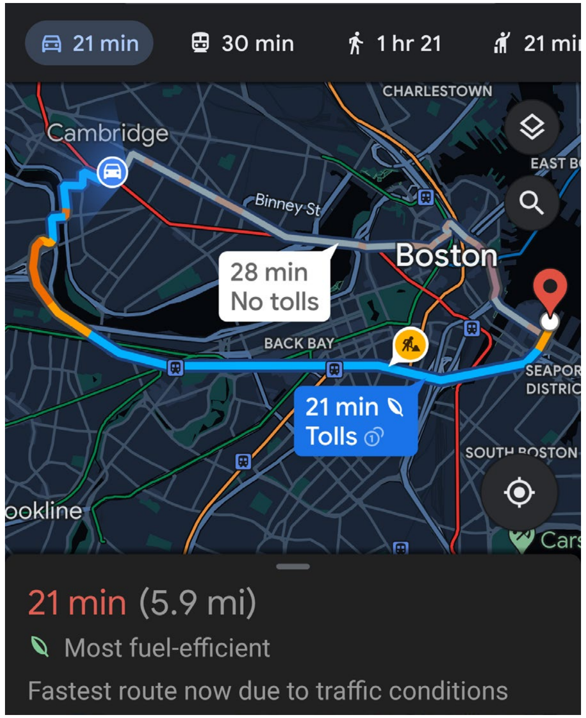


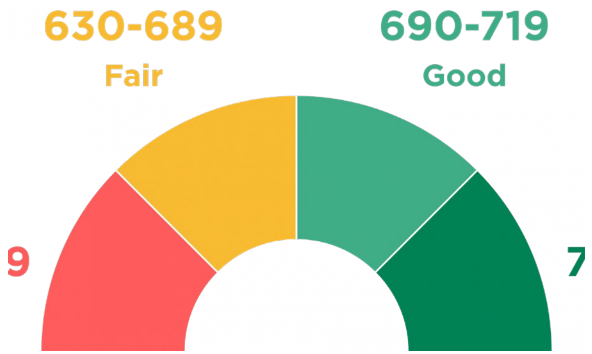
Connections

Traffic prediction

Credit score

wildlife poaching predictions





More drivers follow=> large λ=> chaos

More banks=> large n=> chaos

More dependency => large λ=> chaos

ML in the field

- Multi agent performative prediction
  - How can we learn and avoid chaos?
    - Algorithm
    - Mechanism
  - Competition between learning agents
- When are predictions performative?
  - Strategic classification
  - Price competition