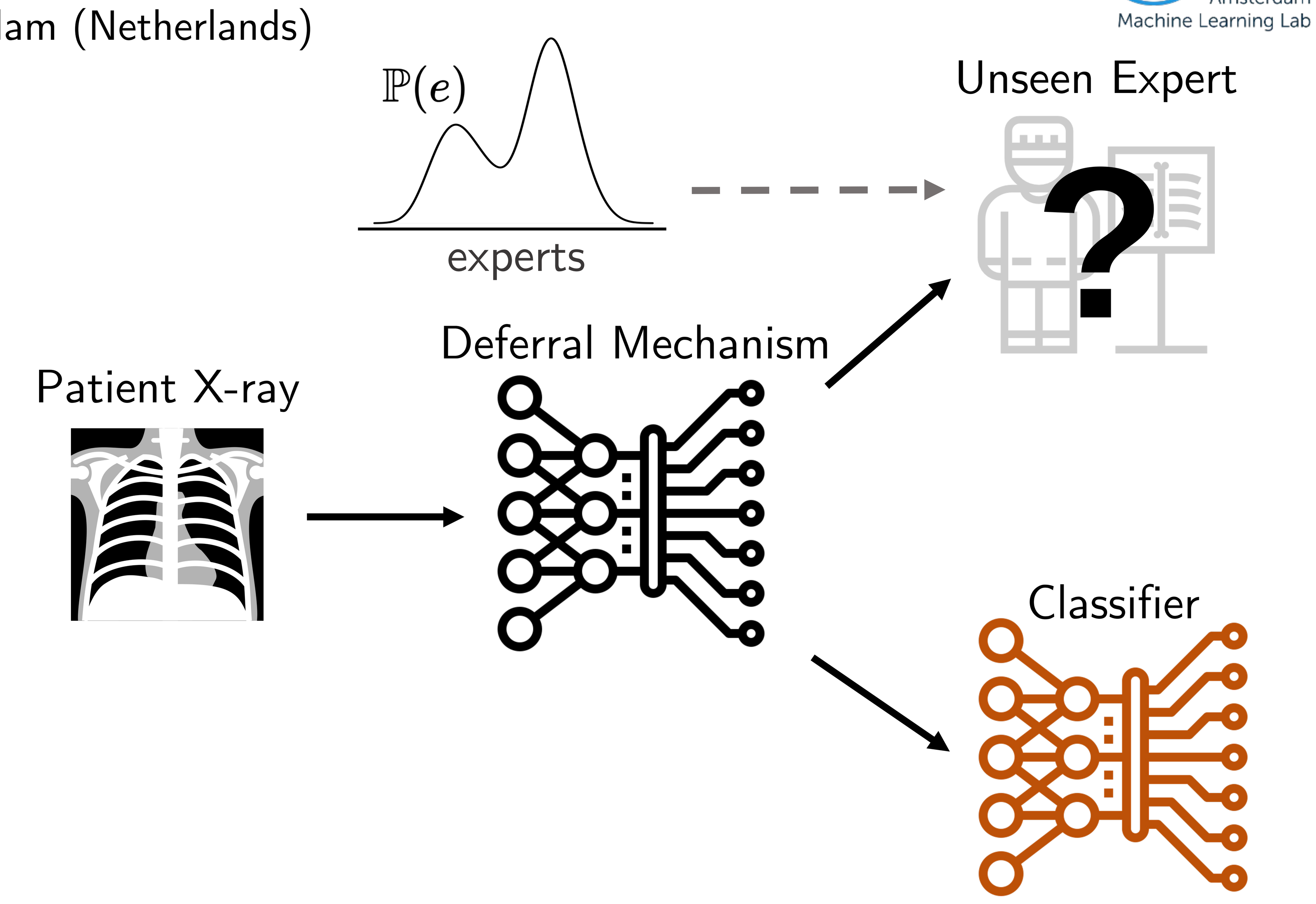
Learning to Defer to a Population: A Meta-Learning Approach

Dharmesh Tailor, Aditya Patra, Rajeev Verma, Putra Manggala, Eric Nalisnick
University of Amsterdam (Netherlands)

Learning-to-defer (L2D) is a proposal for **hybrid intelligent** systems that gives the AI the option to abstain and defer its prediction to a human upon facing a challenging or high-risk decision. Existing L2D systems are trained to be customized to one (or more) specific humans and if the expert were to change, the system should be re-trained.

We formulate an L2D system that can cope with **unseen experts** at test-time by training its deferral subcomponent to generalize to all experts in a population. We propose a general **meta-learning** implementation that can adapt to any expert by only using a **context set** of demonstrations.



Background: L2D [Mozannar & Sontag, 2020]

The learning problem involves jointly training two sub-models: a classifier $h : \mathcal{X} \to \mathcal{Y}$ and a rejector $r : \mathcal{X} \to \{0, 1\}$. When r(x) = 0, the classifier makes the decision, and when r(x) = 1 the classifier abstains and defers the decision to the human.

$$\mathcal{D} = \{\mathbf{x}_n, y_n, m_n\}_{n=1}^N$$

Classifier-rejector (0-1) loss:

$$L_{0-1}(h, r) = \mathbb{E}_{\mathrm{x,y,m}}[\underbrace{(1 - r(\mathbf{x}))}_{\text{predict}}\overbrace{\mathbb{I}[h(\mathbf{x}) \neq \mathrm{y}]}^{\text{classifier cost}} + \underbrace{r(\mathbf{x})}_{\text{defer}}\overbrace{\mathbb{I}[\mathrm{m} \neq \mathrm{y}]}^{\text{expert cost}}]$$

They propose a reduction from multiclass expert deferral to cost sensitive learning by unifying the classifier and rejector via an augmented label space that includes the rejection option: $\mathcal{Y}^{\perp} = \mathcal{Y} \cup \{\perp\}$

Then a (consistent) surrogate loss is constructed that extends cross entropy loss. [Verma & Nalisnick, 2022] also showed one-vs-all parameterization is also a consistent surrogate.

$$\phi_{\mathrm{SM}}(g_1, \ldots, g_K, g_{\perp}; \boldsymbol{x}, y, m) = -\log\left(\frac{\exp\{g_y(\boldsymbol{x})\}}{\sum_{y' \in \mathcal{Y}^{\perp}} \exp\{g_{y'}(\boldsymbol{x})\}}\right)$$

Rejection function:

$$\mathbb{I}[g_{\perp}(\mathbf{x}) \geq \max_k g_k(\mathbf{x})]$$

$$-\,\mathbb{I}[m = y]\ \log\left(\frac{\exp\{g_{\perp}(\boldsymbol{x})\}}{\sum_{y' \in \mathcal{Y}^{\perp}} \exp\{g_{y'}(\boldsymbol{x})\}}\right)$$

Learning to Defer to a Population "L2D-Pop"

We assume a generative process for experts from which experts can be sampled indefinitely and without repetition.

$$\mathfrak{E} \sim \mathbb{P}(\mathfrak{E}), \qquad \mathrm{m} \sim \mathbb{P}(\mathrm{m}|\mathbf{x}, \mathrm{y}, \mathfrak{E})$$

Formulation resembles single-expert L2D but now rejector also takes as input some representation of the currently-available expert.

$$r : \mathcal{X} \times \mathfrak{E} \to \{0, 1\}$$

Surrogate loss:

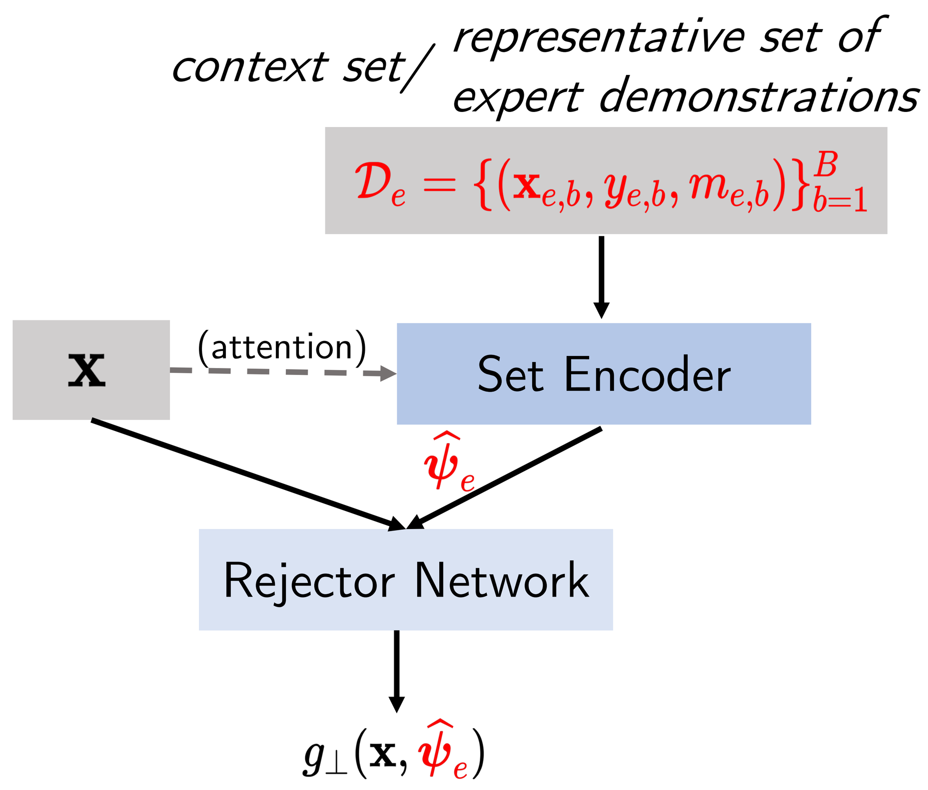
$$\phi_{\text{SM-Pop}}\left(g_1, \ldots, g_K, g_{\perp}; \boldsymbol{x}, y, \{m_e, \psi_e^{\mathfrak{E}}\}_{e=1}^E\right) = \sum_{e=1}^{E} -\log\left(\frac{\exp\{g_y(\boldsymbol{x})\}}{\mathcal{Z}(\boldsymbol{x}, \psi_e^{\mathfrak{E}})}\right) - \mathbb{I}[m_e = y]\ \log\left(\frac{\exp\{g_{\perp}(\boldsymbol{x}, \psi_e^{\mathfrak{E}})\}}{\mathcal{Z}(\boldsymbol{x}, \psi_e^{\mathfrak{E}})}\right)$$

Applying Single-Expert L2D to L2D-Pop:

We can apply single-expert L2D to the population setting to learn a rejector that models the population's marginal probability of correctness. This requires a reformulation of L2D-Pop surrogate loss where rejector no longer a function of the expert and the sum over experts 'pushes through' to the indicator term.

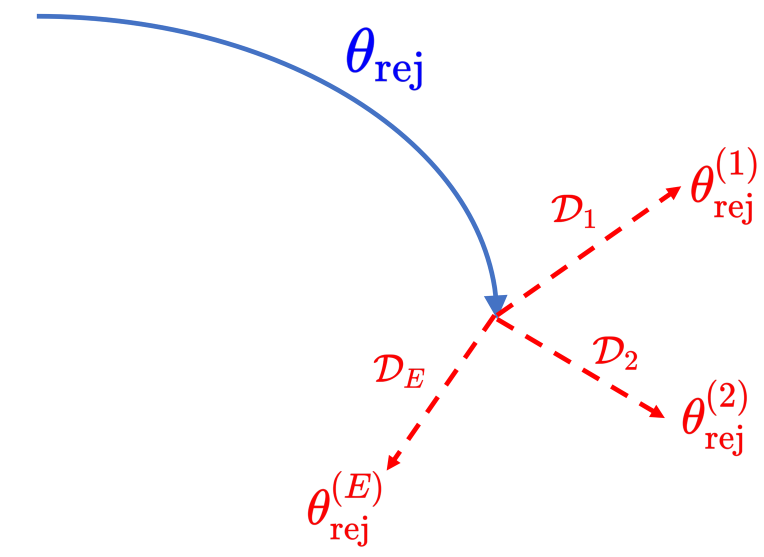
Meta-Learning to Defer

Model-based approach



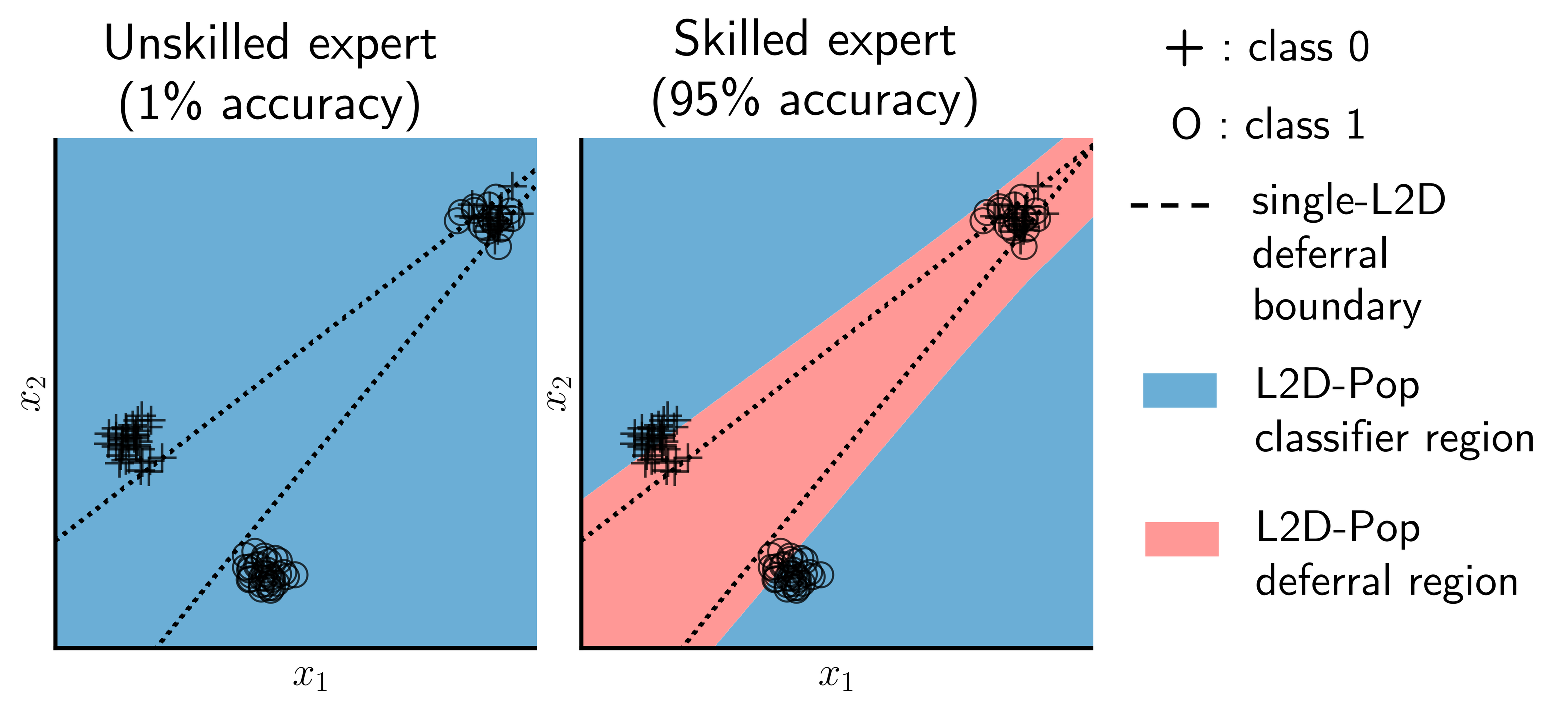
Meta-optimization approach

1. Model the *marginal* expert by the single-expert formulation
2. Then finetune on expert context set at test-time

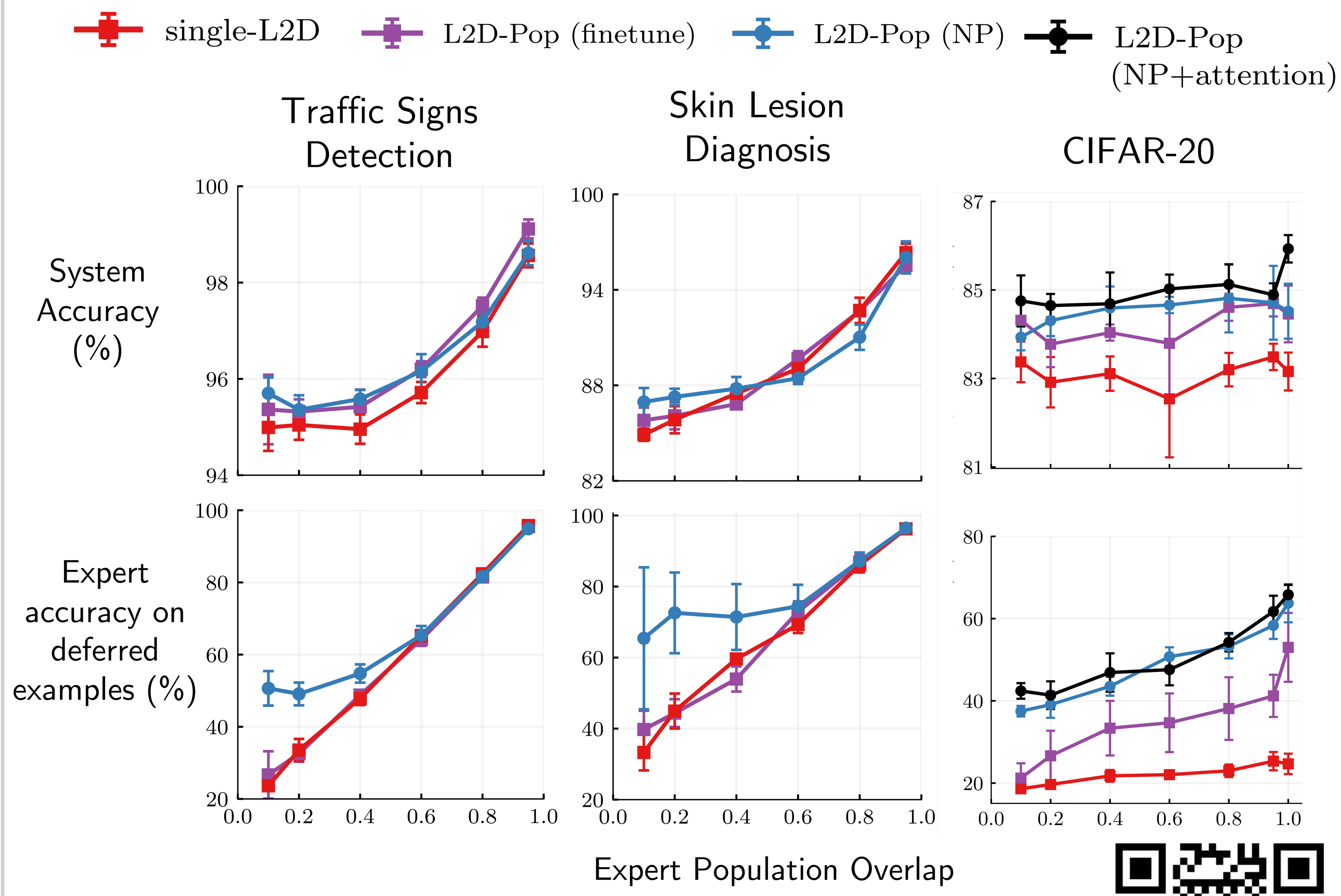


Experiments

Synthetic Data



Varying Population Diversity



References

Mozannar and Sontag. *Consistent Estimators for Learning to Defer to an Expert.* ICML, 2020.

Verma and Nalisnick. *Calibrated Learning to Defer with One-vs-All Classifiers.* ICML, 2022.