Defining Expertise

Applications to Treatment Effect Estimation

Alihan Hüyük,* Qiyao Wei,* Alicia Curth, Mihaela van der Schaar

van_der_Schaar \ LAB
vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

qw281@cam.ac.uk
linkedin.com/in/qiyaowei

Expertise and Treatment Effect Estimation

Decision-makers are often experts of their domains

(often the main cause of confounding)

"Expertise" can be informative in TE estimation!

A TOY EXAMPLE

- **Outcomes:** $Y_0 = 0, Y_1 \sim \mathcal{N}(T, 1)$ *where* T *is the ATE*
- **Prior:** $T \sim \mathcal{N}(0,1)$
- **Suppose** $A \sim \mathcal{B}(0.5)$: $\mathbb{E}[T|Y_1 = y_1] = y_1$
- **Suppose** $A = \mathbb{I}\{T > 0\}$: $\mathbb{E}[T|Y_1 = y_1, A = 1] > y_1$

Yet, the overlap assumption is usually the only assumption made

(expertise is ignored as an information source)

Definitions of Expertise

Predictive Expertise	Prognostic Expertise
$A \sim \pi(X) = \pi(\mu_1(X) - \mu_0(X))$	$A \sim \pi(X) = \pi(\mu_0(X), \mu_1(X))$
e.g. doctors try to assign the treatments with the largest benefit to the patient	*e.g.* teachers might target students who struggle the most (not who would benefit the most)
$E^{\pi}_{\text{pred}} = \frac{I(A^{\pi}; Y_1 - Y_0)}{\mathbb{H}[A^{\pi}]}$	$E^{\pi}_{\text{prog}} = \frac{I(A^{\pi}; Y_0, Y_1)}{\mathbb{H}[A^{\pi}]}$

Different Types of Expertise

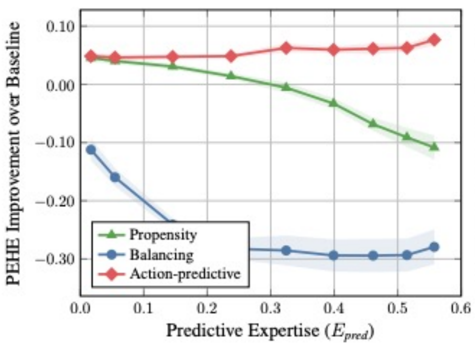
Predictive Expertise	Prognostic Expertise
$A \sim \pi(X) = \pi(\mu_1(X) - \mu_0(X))$	$A \sim \pi(X) = \pi(\mu_0(X), \mu_1(X))$
e.g. doctors try to assign the treatments with the largest benefit to the patient	*e.g.* teachers might target students who struggle the most (not who would benefit the most)

Knowing the type of expertise prominent in a domain is important!

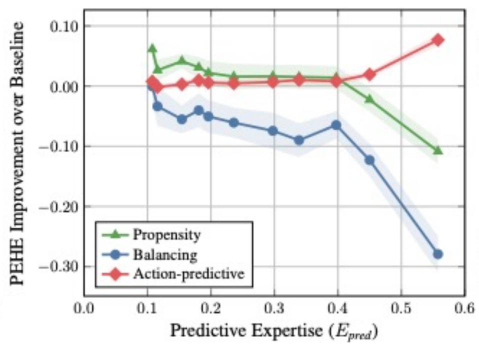
Consider balancing representations:

- *learns* $Z = \rho(X)$ *such that* $A \perp Z$
- *would not work well under predictive expertise*
- *might work well under prognostic expertise (w/o predictive expertise)*

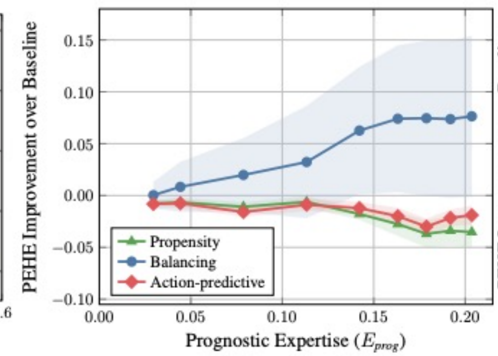
Performance under Different Expertise Scenarios



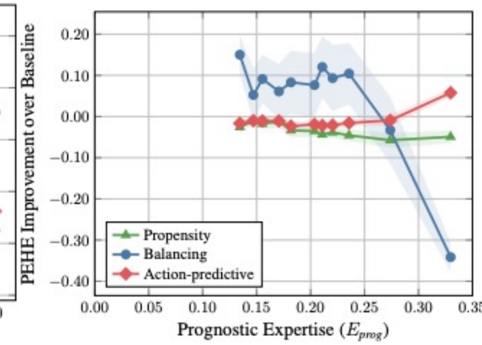
(a) Best→Expert for predictive expertise



(b) Worst→Expert for predictive expertise



(c) Best→Expert for prognostic expertise

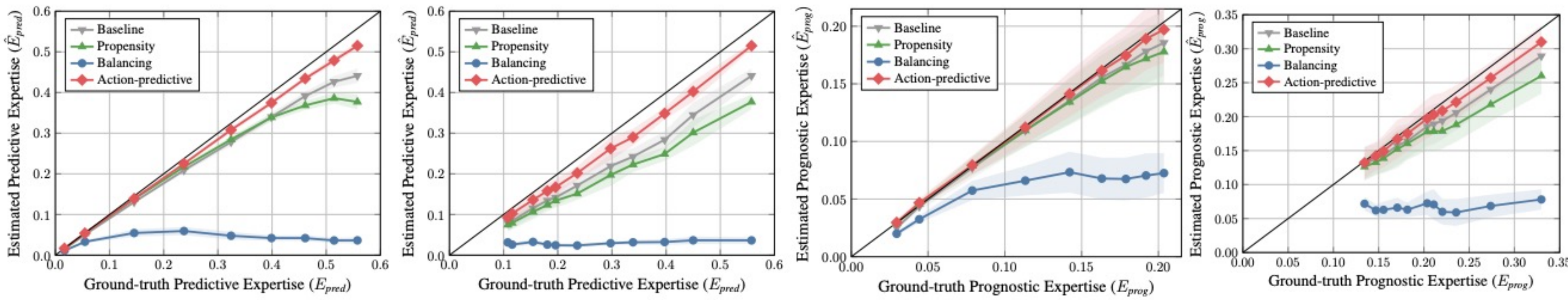


(d) Worst→Expert for prognostic expertise

Predictive Expertise **Prognostic Expertise**

- On average, *Action-Predictive* performs the best under *Predictive Expertise*
  Balancing performs the best under *Prognostic Expertise*

Estimating Expertise

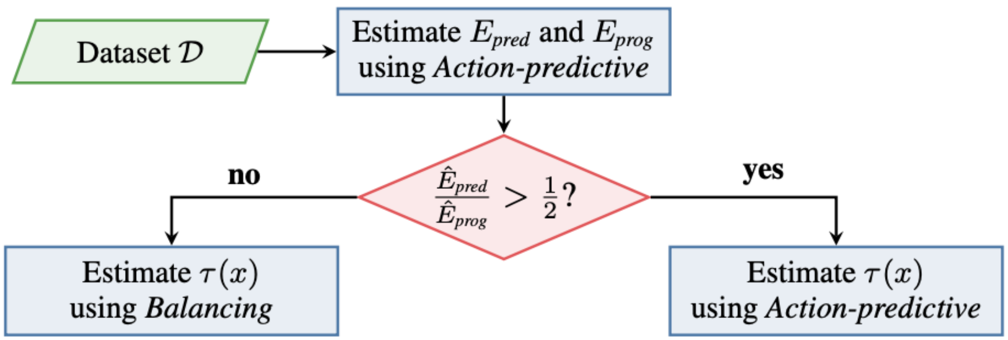


(a) Best→Expert for predictive expertise *(b) Worst→Expert for predictive expertise* *(c) Best→Expert for prognostic expertise* *(d) Worst→Expert for prognostic expertise*

Predictive Expertise **Prognostic Expertise**

- However, *Action-Predictive* is uniformly better at predicting expertise!

Expertise-Informed Treatment Effect Estimation



Method	**Predictive Datasets**	**Prognostic Datasets**	**All Datasets**
Baseline	0.784 (0.130)	1.483 (0.258)	1.134 (0.194)
Propensity	0.786 (0.126)	1.511 (0.259)	1.149 (0.193)
Balancing	0.936 (0.131)	**1.439 (0.243)**	1.188 (0.187)
Action-predictive	**0.751 (0.128)**	1.495 (0.259)	1.123 (0.194)
Expertise-informed	**0.751 (0.128)**	**1.439 (0.243)**	**1.096 (0.185)**

(Best-of-Both-Worlds)

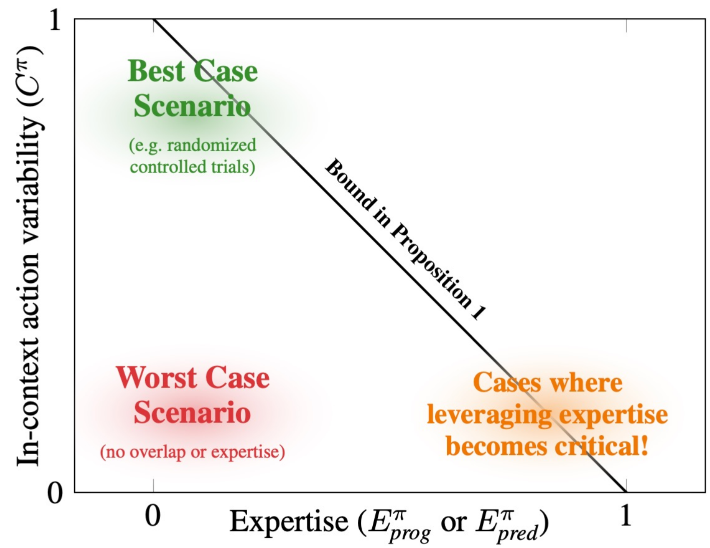
More details in our paper!

Defining Expertise:
Application to Treatment Effect Estimation

Theory of Expertise

Predictive Expertise	Prognostic Expertise
$E^{\pi}_{\text{pred}} = I(A^{\pi}; Y_1 - Y_0)/\mathbb{H}[A^{\pi}]$	$E^{\pi}_{\text{prog}} = I(A^{\pi}; Y_0, Y_1)/\mathbb{H}[A^{\pi}]$

- **In-context Action Variability (i.e. Overlap):** $C^{\pi} = \mathbb{H}[A^{\pi}|X]/\mathbb{H}[A^{\pi}]$
- **Proposition 1:** $E^{\pi} + C^{\pi} \leq 1$



Theory of Expertise: Experimental Setup