1 Institute of Science and Technology Austria

2 Inria

NEURAL INFORMATION PROCESSING SYSTEMS

Smoke and Mirrors in Causal Downstream Tasks

Riccardo Cadei[1]
Lukas Lindorfer[1]
Sylvia Cremer[1]
Cordelia Schmid[2]
Francesco Locatello[1]

Causal Learning and AI

Motivation: *Experimental Ecology*

Treated Group
Batch: 1, Position: 6, Time: 2m45s

X

Y

No Action

Objective:

$$ATE := \mathbb{E}[Y|do(T=1)] - \mathbb{E}[Y|do(T=0)]$$

Treatment
(e.g., virus)

Behaviour
(e.g., grooming)

Motivation: *Experimental Ecology*

Treated Group
Batch: 1, Position: 6, Time: 2m45s

X

f

$\hat{Y}$

No Action (99.93%)

Objective:

$$ATE \approx \mathbb{E}[f(\boldsymbol{X})|do(T=1)] - \mathbb{E}[f(\boldsymbol{X})|do(T=0)]$$

Treatment
(e.g., virus)

Predicted Behaviour
(e.g., grooming)

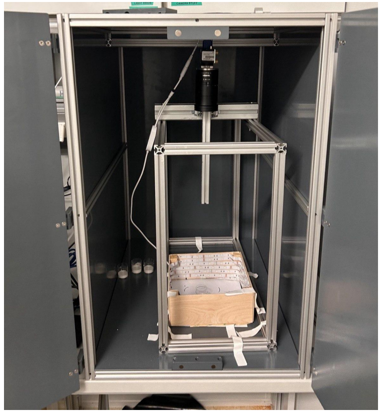
Treatment Effect Bias

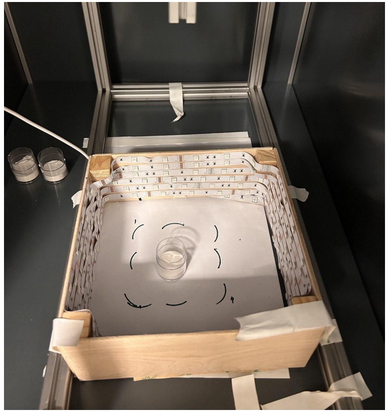
$$TEB := \Big(\underbrace{\mathbb{E}_{\boldsymbol{X}|do(T=1)}[f(\boldsymbol{X})] - \mathbb{E}_{Y|do(T=1)}[Y]}_{\textit{Interventional Bias under Treatment}}\Big) - \Big(\underbrace{\mathbb{E}_{\boldsymbol{X}|do(T=0)}[f(\boldsymbol{X})] - \mathbb{E}_{Y|do(T=0)}[Y]}_{\textit{Interventional Bias under Control}}\Big)$$

Contributions

- Formulation challenges in (Factual) Estimation for Causal Downstream Tasks
  - Potential mitigations
  - Experimental evidence (synthetic and real-world data)

- [New Benchmark] ISTAnt*

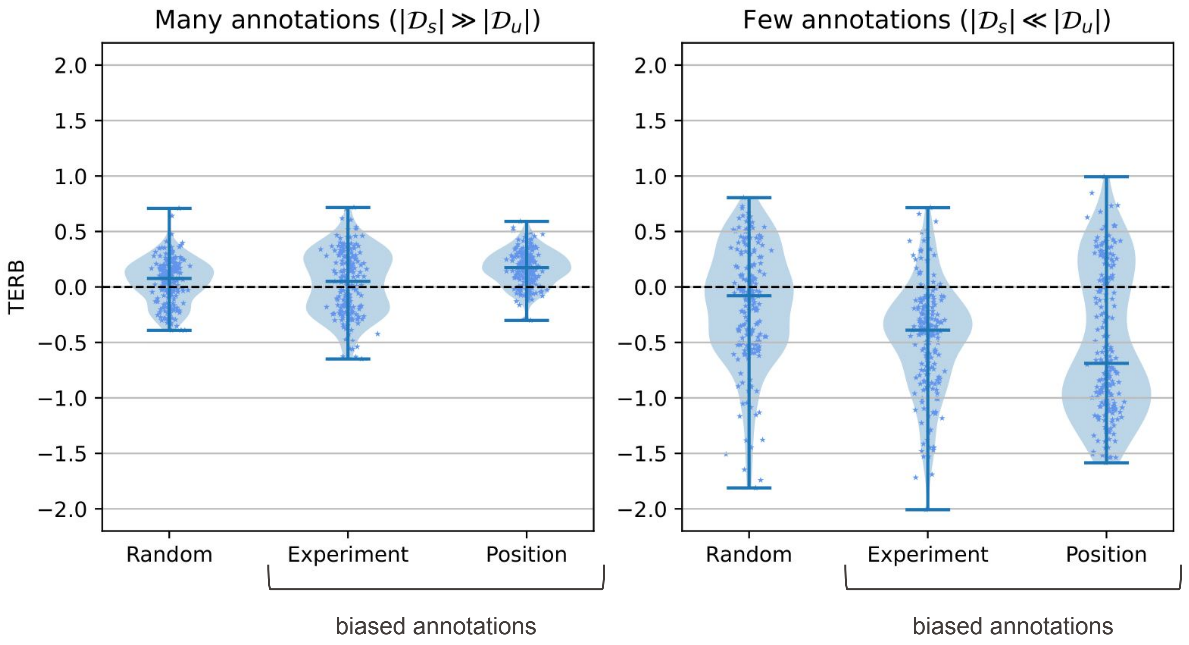




*first benchmark for Representation Learning for Causal Downstream Tasks

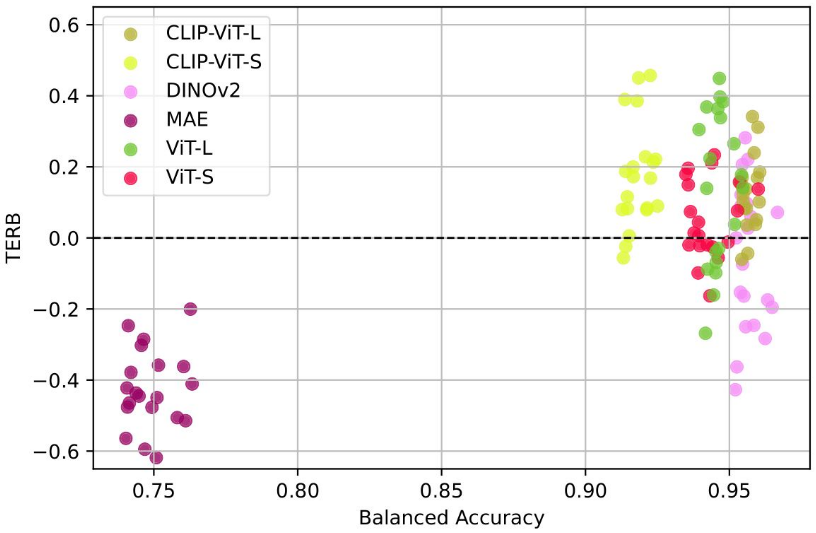
Issue 1: *Annotation Bias*

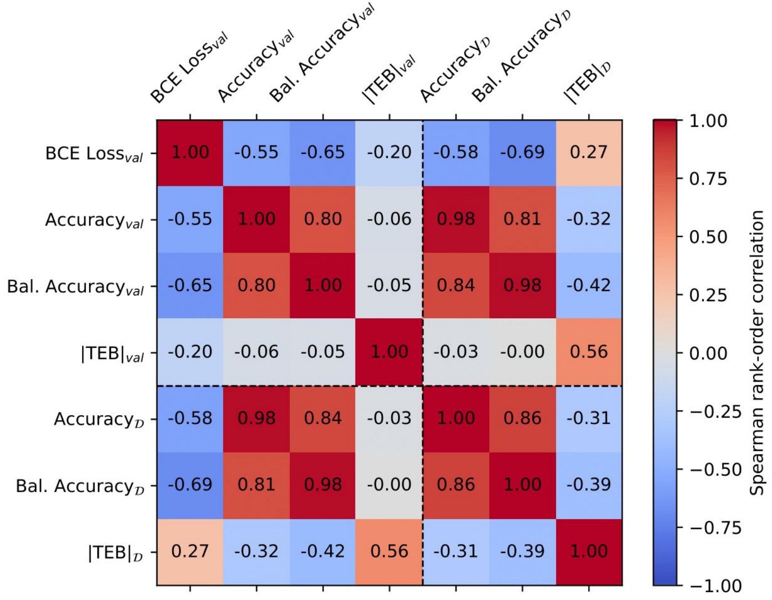
Unbiased supervision (i.e., annotation) is crucial to close backdoor paths and avoid confounding effects



Issue 2: *Encoder Bias*

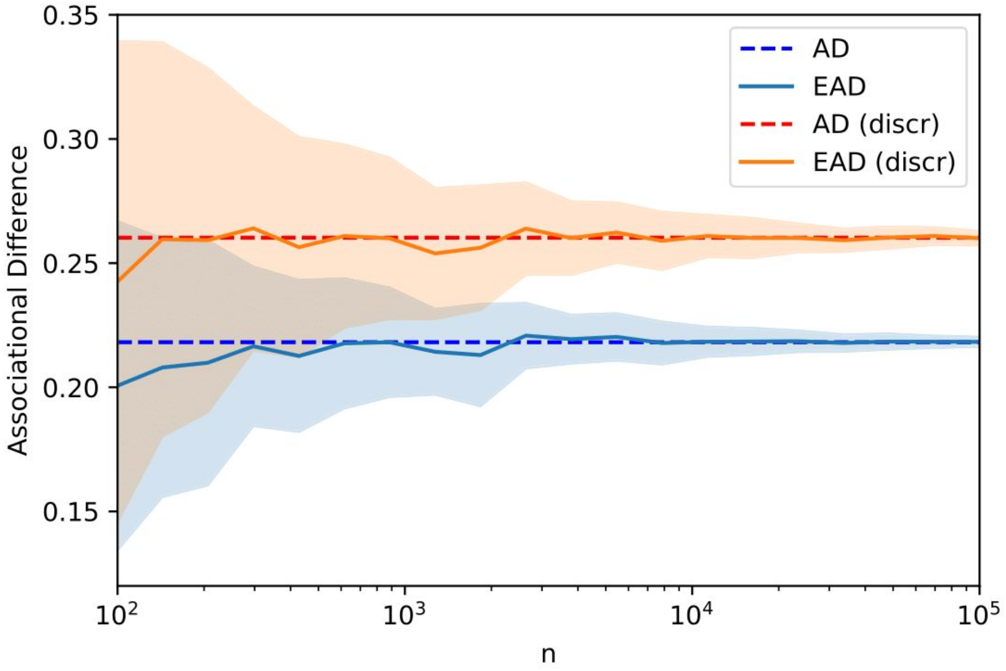
Prediction is not Causal Estimation. New methodologies to mitigate the treatment effect bias during adaptation should be investigated.





Issue 3: *Discretization Bias*

Don't arbitrarily discretize predictions for downstream treatment effect estimation.



Reproducibility

Code:

https://github.com/CausalLearningAI/ISTAnt
https://github.com/CausalLearningAI/CausalMNIST

Data:

https://doi.org/10.6084/m9.figshare.26484934.v2

Thank you for your attention