Learning Counterfactual Representations for Estimating Individual Dose-Response Curves

Patrick Schwab[1] (@schwabpa), Lorenz Linhardt[2], Stefan Bauer[3], Joachim M. Buhmann[2] and Walter Karlen[1]

[1] Institute of Robotics and Intelligent Systems, ETH Zurich, Switzerland
[2] Department of Computer Science, ETH Zurich, Switzerland
[3] MPI for Intelligent Systems, Tübingen, Germany

1 Introduction

Estimating what would be an **individual's potential response to varying levels of exposure to a treatment** is of high practical relevance for several important fields, such as healthcare, economics and public policy.

How can we train models to estimate counterfactual outcomes in settings **with any number of treatments and associated dosage parameters**?

2 Dose-response Networks (DRNet)

The core idea behind our approach is to leverage **conditional computation** in **Dose-Response Networks** (DRNets). To maintain a strong influence of treatment and dosage on the final prediction, DRNets provide **independent prediction paths** through a neural network for **each treatment** and for a **configurable number *E* of strata** of the range of possible dosage values. **Parameter sharing** across treatments and across dosage strata is handled through a **hierarchy of shared base layers.** The model is trained end-to-end on observed samples.

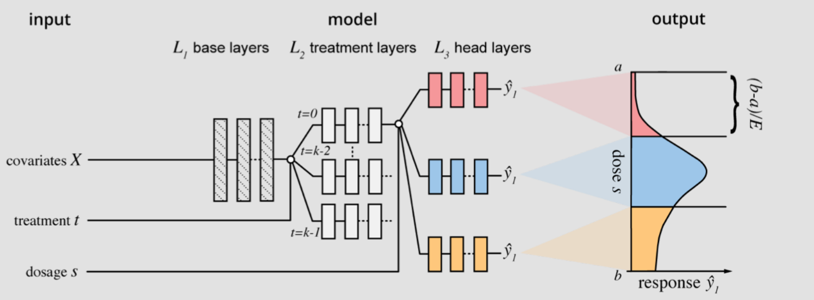
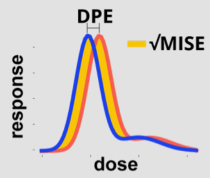


Figure 1. An overview of Dose-Response Networks (DRNets) with shared base layers, per-treatment layers and head networks for a configurable number *E* of dosage strata.

3 Metrics, regularisation and benchmarks

Estimating dose-response to multiple treatments with dosages poses challenges in terms of **model evaluation, model regularization,** and **method benchmarking.** To address these challenges, we introduce:

Performance metrics for estimating individual dose-response:
Mean Integrated Square Error (MISE), Dosage Policy Error (DPE)



Open benchmarks for counterfactual inference of dose-response:

News. Simulates reader's opinion on exposure to articles on different viewing devices. (n=5000, p=2870 features, T=2/4/8/16 treatments)

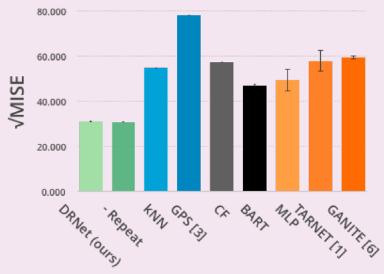
Mechanical Ventilation in the Intensive Care Unit (MVICU). Simulates patient response to mechanical ventilation. (n=8040, p=49, T=3)

The Cancer Genomic Atlas (TCGA). Simulates individual response to medication, surgery and chemotherapy. (n=9659, p=20531, T=3)

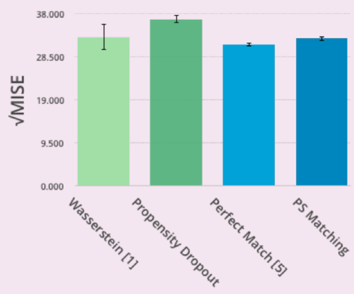
4 Results

DRNets **outperform** existing methods **across all three benchmarks**.

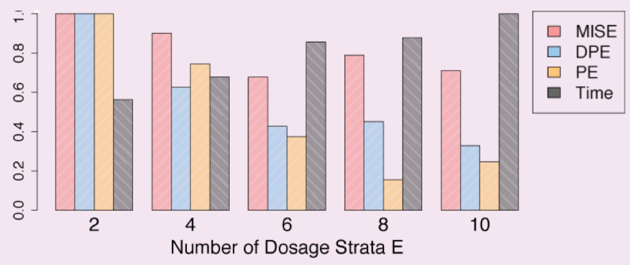
MVICU



Across benchmarks, Wasserstein [1] and Perfect Match [5] were slightly more effective than Propensity Score Matching and Propensity Dropout.



Tradeoff between computational and predictive performance: Experimentally, a higher number of dosage strata was associated with increases in predictive performance and computation time.



5 Conclusion

We present …

- an approach for estimating individual dose-response that works with **multiple treatment options with associated dosage parameters**
- extensions of several existing **regularization methods** for this setting
- **performance metrics** for evaluating dose-response estimators
- **open benchmarks** for comparing dose-response estimators

Source code is available at **https://github.com/d909b/drnet**

6 References

1. Shalit, U., Johansson, F.D., and Sontag, D. 2017. Estimating Individual Treatment Effect: Generalization Bounds and Algorithms. International Conference on Machine Learning.
2. Rosenbaum, P.R. and Rubin, D.B.. 1983. The Central Role of the Propensity Score in Observational Studies of Causal Effects. Biometrika 70(1):41-55
3. Imbens, G.W. 2000. The Role of the Propensity Score in Estimating Dose-Response Functions. Biometrika, 87(3):706-710
4. Schuler, A., Baiocchi, M., Tibshirani, R. and Shah, N. 2018. A Comparison of Methods for Model Selection when Estimating Individual Treatment Effects. arXiv preprint arXiv:1804.05146
5. Schwab, P., Linhardt, L., and Karlen, W. 2018. Perfect Match: A Simple Method for Learning Counterfactual Representations with Neural Networks. arXiv preprint 1810.00656
6. Yoon, J., Jordon, J. and van der Schaar, M. 2018. GANITE: Estimation of Individualized Treatment Effects using Generative Adversarial Nets. International Conference on Learning Representations.

AAAI | 2020

Swiss National Science Foundation

75 NRP Big Data National Research Programme