CausalBench: Causal Learning Research Streamlined

Ahmet Kapkiç, Pratanu Mandal, Abhinav Gorantla, Shu Wan, Ertuğrul Çoban, Paras Sheth
Huan Liu, K. Selçuk Candan
{akapkic,pmandal5,agorant2,swan16,ecoban,psheth5,huanliu,candan}@asu.edu
Arizona State University, Tempe, AZ

Abstract

Recent advances in causal machine learning introduced a plethora of new causal discovery and causal inference models to tackle decision support problems. Yet, these models exhibit different performance when they train on different data, and even different hardware/software platforms, making it challenging for users to select the appropriate setup pertinent to their specific problem instance. The situation is complicated by the fact that, until recently, the field lacked a unified, publicly available, and configurable platform that supports all major causal inference tasks, including causal discovery, causal effect estimation, and causal inference. CausalBench[1] is a comprehensive benchmarking tool for causal machine learning that facilitates accurate and reproducible benchmarking of causal models across metrics and deployment contexts and helps users to select the most appropriate set up (such as hyper-parameter configuration) for the specific problem setting. This tutorial is intended to familiarize attendees from diverse backgrounds, who are interested in causal learning models and with the capabilities of CausalBench. The tutorial begins with an introduction to "causality" and causal machine learning, and then provides hands-on experience with CausalBench to equip attendees with the knowledge necessary to utilize CausalBench for their causal learning problems.

Keywords

Causality, Causal Learning, Benchmarking,

ACM Reference Format:
Ahmet Kapkiç, Pratanu Mandal, Abhinav Gorantla, Shu Wan, Ertuğrul Çoban, Paras Sheth, Huan Liu, K. Selçuk Candan. 2025. CausalBench: Causal Learning Research Streamlined. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2 (KDD '25), August 3–7, 2025, Toronto, ON, Canada.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3711896.3737598

1 Introduction

One of the critical challenges faced by the causal machine learning research community is reproducibility of experiments. Even with the current best efforts to provide detailed experiment setups, a change in a single driver or library can cause significant differences in the results. An ideal benchmarking platform would document every aspect of an experiment, from the data to the hardware/software configuration of the system used for running the experiments.

CausalBench [4] is a platform for publicly available benchmarks and consensus-building standards for the evaluation of causal learning models and algorithms from observational data[2]. It aims to assist researchers and developers in easily applying and effectively evaluating (a) causal inference, (b) causal discovery, and (c) causal interpretability algorithms with a variety of standard metrics, procedures, and large-scale datasets. CausalBench provides a standard and convenient way for the community to contribute data, models, and metrics (Figure 1). Moreover, for the community to trust the results included in the benchmark, CausalBench provides a transparent mechanism to track and log an experiment, whether on the data, the model, or the experiment context itself. When comparing experimental results, CausalBench helps identify differences in experimental setups that could potentially explain the outcomes and recommends new experiment set ups.

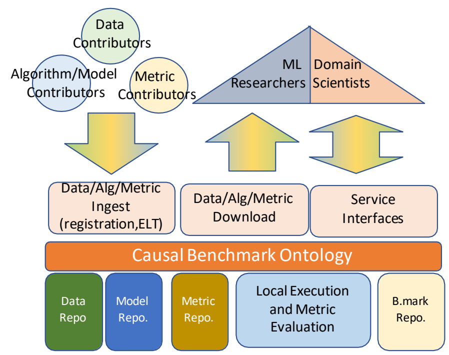

Figure 1: Overview of CausalBench (causalbench.org)

2 Focus Group

This tutorial is designed for students, researchers, and practitioners from all backgrounds who are interested in causal learning models. It is especially beneficial for researchers in the ML community who are new to causal learning models and looking for a lecture to quickly step into the field or for experts in causal learning who are looking for a streamlined benchmarking suite with available datasets. We expect the audience to have a general understanding of machine learning, but no prior knowledge of causality or benchmarking is necessary.

[1]This research is funded by NSF Grant 2311716, "CausalBench: A Cyberinfrastructure for Causal-Learning Benchmarking for Efficacy, Reproducibility, and Scientific Collaboration", and NSF Grants 2230748, "PIRE: Building Decarbonization via AI-empowered District Heat Pump Systems", 2412115, "PIPP Phase II: Analysis and Prediction of Pandemic Expansion (APPEX)" and USACE GR40695, "Designing nature to enhance resilience of built infrastructure in western US landscapes".

[2]Documentation and a video explaining how to use CausalBench are available at https://docs.causalbench.org.

KDD '25, Toronto, ON, Canada

ACM ISBN 979-8-4007-1454-2/2025/08
https://doi.org/10.1145/3711896.3737598

3 Outline, Program, and Participation Strategies

This **system-focused**, hands-on tutorial is set around the CausalBench platform, and how to utilize it to streamlining causal machine learning research. The 3-hour long tutorial program (including a 10-minute break) will include the following content:

(1) **Causality: What, Why, How *(1h)***
   - Introduction to causality [6, 7]
   - Models for representing causal knowledge[3, 5]
   - State-of-the art in causal discovery and inference [1, 2, 8]

(2) **CausalBench[4]: Create, Compare, Contribute *(1h 35m)***
   - Introduction to CausalBench
   - Installing CausalBench and setting up a benchmark
   - Benchmark contexts: details and configuration files
   - Creating and executing a benchmark context
   - Analyzing and interpreting the results

(3) **Conclusions *(15m)***
   - Did CausalBench help improve research workflows?
   - Challenges ahead and contributing to CausalBench

During the first hour of the program, the audience will be introduced to the fundamentals of causality and existing work in the field. Intuitive examples and visual representations will be employed to ensure the tutorial is both engaging and easy to understand. Once the basics are covered, the audience will be guided through an interactive tour of the CausalBench framework. We will provide a partially pre-filled Google Colab notebook, allowing participants to follow along and perform hands-on tasks as we demonstrate key features of CausalBench. During this period, attendees will learn how to utilize CausalBench by accessing data sets, models, and metrics, creating and executing benchmark scenarios, and comparing and analyzing benchmarks results.

By the end of the tutorial, the users are expected to have a working knowledge of the CausalBench suite and the its functionalities that will help streamlining their causal machine learning work.

3.1 Resource Requirements

We will use Google Colab for all demonstrations and audience tasks, offering a standardized execution platform with minimal setup overhead. Google Colab will handle the primary computational load of executing the causal models, allowing the audience to use any low-end hardware to complete the tasks. Our demonstration will require a stable internet connection, though the bandwidth requirements will be minimal, as the actual data transfer will occur between Google Colab and the CausalBench servers.

4 Related Past Tutorials

- "Using OpenML for haring datasets, algorithms, and experiments" **Presenter(s)**: Joaquin Vanschoren, Morris Riedel. **Conference**: RAISE CoE Training, Virtual, 2022 **Relationship**: Both tutorials discuss open source, online collaborative benchmarking platforms. **Differences**: The tutorial focuses on AutoML system's API support for popular machine learning tasks. Our tutorial, on the other hand, focuses on benchmarking of causal learning tasks and causal-ML oriented tools for experiment optimization.
- "Tutorial on Causal Inference and Counterfactual Reasoning" **Presenter(s)**: Amit Sharma, Emre Kiciman. **Conference**: ACM KDD 2018 **Relationship**: Both tutorials discuss causal inference machine learning algorithms **Differences**: The tutorial focuses on causal machine learning algorithms and their use cases in real-world problems. In contrast, our tutorial provides hands-on experience with the benchmarking of causal ML algorithms.

5 Tutors and In-Person Presenters

Ahmet Kapkiç[3]**, Pratanu Mandal**[3]**, Abhinav Gorantla, Shu Wan, Ertuğrul Çoban, Paras Sheth** are graduate students in Computer Science at Arizona State University. Their research interests lie at understanding causal dynamics across data, model, and system characteristics within a causal context, and utilizing this knowledge to enhance machine learning approaches.
Dr. Huan Liu is a Professor of Computer Science and Engineering at Arizona State University. His research interests are in data mining, machine learning, social computing, and artificial intelligence. He is Editor in Chief of ACM TIST, and Field Chief Editor of Frontiers in Big Data and its Specialty Chief Editor of Data Mining and Management. He is a Fellow of the ACM, AAAI, AAAS, and IEEE.
Dr. K. Selçuk Candan[3] is a Professor of Computer Science and Engineering at Arizona State University (ASU) and the director of ASU's Center for Assured and Scalable Data Engineering (CASCADE). His primary research interest is in the area of management and analysis of non-traditional, heterogeneous, and imprecise (such as multimedia, web, and scientific) data.

The contact author, Kapkiç, is reachable at *akapkic@asu.edu.*

5.1 Conclusion and Societal Impacts

In critical areas, such as public health and sustainability, grasping the underlying causal connections between actions and consequences is vital. Unfortunately, traditional methods like randomized controlled trials are often impractical or unethical in such contexts. Fortunately, the availability of extensive observational data enables the approximation of causal relationships through data analytics, facilitating the discovery of meaningful patterns and informing effective decision-making. Causal learning from observational data offers a promising alternative to correlation-based learning. This hands-on tutorial will help researchers and practitioners in better optimizing their causal learning workflows.

References

[1] Azad, F. T., Candan, K. S., Kapkiç, A., Li, M.-L., Liu, H., Mandal, P., Sheth, P., Arslan, B., Chowell-Puente, G., Sabo, J., et al. (vision paper) a vision for spatio-causal situation awareness, forecasting, and planning. *ACM Transactions on Spatial Algorithms and Systems 10*, 2 (2024), 1–42.
[2] Guo, R., Cheng, L., Li, J., Hahn, P. R., and Liu, H. A survey of learning causality with data: Problems and methods. *ACM Computing Surveys (CSUR) 53*, 4 (2020), 1–37.
[3] Hoyle, R. H. *Handbook of structural equation modeling.* Guilford press, 2012.
[4] Kapkiç, A., Mandal, P., Wan, S., Sheth, P., Gorantla, A., Choi, Y., Liu, H., and Candan, K. S. Introducing causalbench: A flexible benchmark framework for causal analysis and machine learning. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (2024), pp. 5220–5224.
[5] Pearl, J. Graphs, causality, and structural equation models. *Sociological Methods & Research 27*, 2 (1998), 226–284.
[6] Pearl, J. *Causality.* Cambridge university press, 2009.
[7] Pearl, J., Glymour, M., and Jewell, N. P. *Causal inference in statistics: A primer.* John Wiley & Sons, 2016.
[8] Wan, S., Shah, R., Deng, Q., Sabo, J., Liu, H., and Candan, K. S. Spatio-temporal causal learning for streamflow forecasting. In *2024 IEEE International Conference on Big Data (BigData)* (2024), IEEE, pp. 6161–6170.

[3]Designated authors are expected to present in-person; however, the final list of presenters will be announced on the tutorial website at tutorial.causalbench.org.