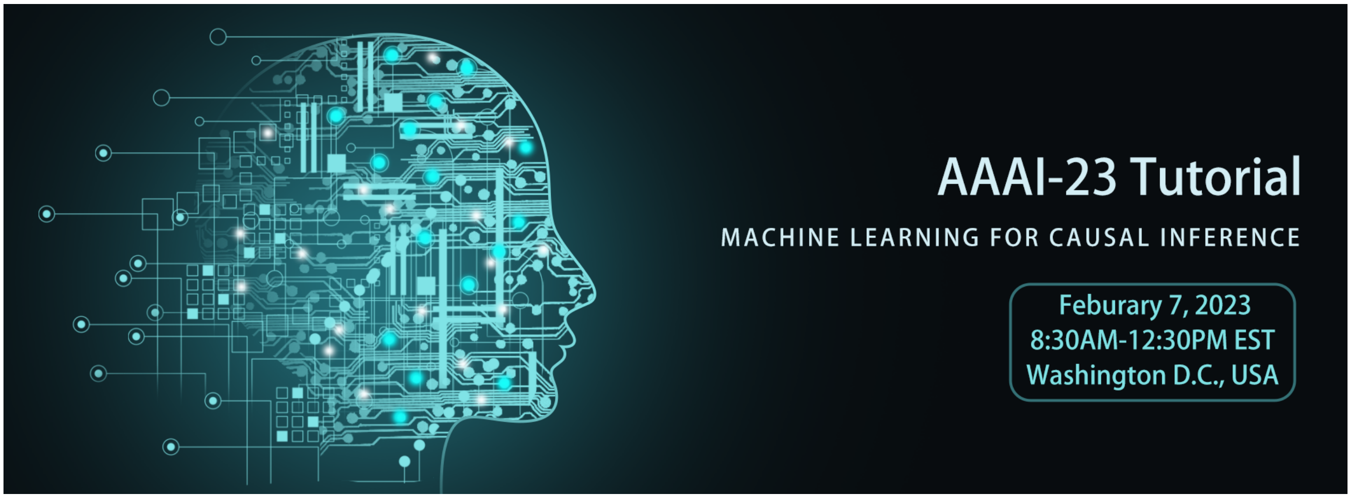

Zhixuan Chu[1], Jing Ma[2], Jundong Li[2], Sheng Li[2]
[1]Ant Group, China
[2]University of Virginia, USA
chuzhixuan.czx@alibaba-inc.com, {jm3mr, jundong, shengli}@virginia.edu

Time: Tuesday, 2/7/23, 8:30 AM - 12:30 PM

Location: Walter E. Washington Convention Center, Washington DC, USA

Presenters:

- Zhixuan Chu, Ant Group, China.
  Email: chuzhixuan.czx@alibaba-inc.com
- Jing Ma, University of Virginia, USA.
  Email: jm3mr@virginia.edu
- Jundong Li, University of Virginia, USA.
  Email: jundong@virginia.edu
- Sheng Li, University of Virginia, USA.
  Email: shengli@virginia.edu

Tutorial Website:
https://aaai23causalinference.github.io/

Tutorial Slide:
https://aaai23causalinference.github.io/AAAI2023MachineLearningforCausalInference.pdf

Tutorial Outline

The causal inference has numerous real-world applications in many domains such as health care, marketing, political science, and online advertising. Treatment effect estimation, a fundamental problem in causal inference, has been extensively studied in statistics for decades. However, traditional treatment effect estimation methods may not well handle large-scale and high-dimensional heterogeneous data. In recent years, an emerging research direction has attracted increasing attention in the broad artificial intelligence field, which combines the advantages of traditional treatment effect estimation approaches (e.g., matching estimators) and advanced machine learning approaches (e.g., adversarial learning, graph neural networks). In this tutorial, we will start with a brief overview of traditional causal inference methods, and then focus on introducing state-of-the-art machine learning algorithms for causal inference, especially for the treatment effect estimation task. In addition, we will showcase promising applications of these methods in multiple application domains.

Schedule

The general outline of this tutorial is listed below:

- Background on Causal Inference
- Representation Learning based Methods
- Graph Neural Networks based Methods
- Causal Inference-aided Machine Learning
- Causal Inference Applications
- Conclusions and Future Perspectives

Content

Background on Causal Inference

We will first motivate the presented topic with real-world examples, such as causal questions raised in education, marketing, health care, political science, and online advertising. Then, we will introduce the basic concepts of causal inference, as well as the fundamental causal inference frameworks, such as the potential outcome framework and causal graphical models. We will also use illustrative examples to explain the concepts and terminology. After that, the formal definitions and implications of several important concepts in causal inference will be introduced, including counterfactuals, average treatment effects (ATE), average treatment effects on treated units (ATT), and individual treatment effects (ITE). Major approaches for causal inference, including both experimental study and observational study will be briefly introduced as well. We will focus on the potential outcome framework for an observational study, and present the problem definitions with essential assumptions.

Representation Learning based Methods

In this part, we will first briefly introduce traditional causal inference methods such as nearest neighbor matching and propensity score matching, discuss some interesting theoretical results, and then motivate the representation learning methods for effective estimation of causal inference. The concepts and foundations of subspace learning will be discussed as well. We will cover both subspace learning and deep representation learning methods.

For subspace learning methods, we will introduce their ideas and discuss their advantages and disadvantages. These algorithms include but are not limited to: (1) an informative subspace learning method for counterfactual learning, which maximizes the nonlinear dependence between the candidate subspace and the outcome variable measured by the Hilbert-Schmidt independence criterion; and (2) a nonlinear and balanced subspace learning algorithm, which converts counterfactual inference problem to a classification problem and learns a kernel space for effective matching.

For deep representation learning methods, we will first explain why causal inference and treatment effect estimation would benefit from deep representation learning, and then introduce the state-of-the-art deep representation learning based methods. These methods include but are not limited to: (1) a balanced representation learning method, which imposes some regularizations related to treatment group distance upon the representation space, to get the balanced representation; (2) deep feature selection based methods for treatment effect estimation with theoretical analysis; (3) adversarial learning based ITE estimation methods, which leverages the Generative Adversarial Nets (GANs) to capture the counterfactual distributions.

Graph Neural Networks based Methods

In this part, we will go beyond traditional causal inference settings on independent and identically distributed (i.i.d.) data and introduce recent progress of causal inference on graphs. Especially, we focus on those graph neural network (GNN) based methods due to their effectiveness in these problems. Specifically, we will cover two types of works: graph-based deconfounding and network interference.

Graph-based deconfounding is motivated by the issue of unobserved confounders. Confounders (variables that influence both treatment and outcome) can lead to biased causal effect estimation if they are not properly handled. Traditional causal inference relies on the unconfoundedness assumption that there do not exist any unobserved confounders. However, this assumption is often violated in the real world. Graph-based deconfounding methods capture the unobserved confounders by leveraging the graph structure among different units. These methods include but are not limited to: (1) a network deconfounder for ITE estimation, which uses GNN to control for the unobserved confounders buried in graph structure; (2) a dynamic network deconfounder, which considers the whole problem in an evolving networked environment with time-varying treatments, outcomes, confounders, and network structure.

Network interference methods relax the stable unit treatment value assumption (SUTVA) in traditional causal inference settings. With the existence of interference, the treatment of a unit can causally influence the outcome of another unit through their connection in a graph. We will introduce the background and the state-of-the-art GNN-based methods for network interference. These methods include but are not limited to: (1) a node representation learning method, which models the interference through graph propagation in GNN; (2) a hypergraph neural network based method which models the high-order interference on hypergraphs.

Causal Inference-aided Machine Learning

In this part, we will introduce how to exploit causality instead of plausible correlations for effectively addressing the robustness and interpretability challenges posed by conventional machine learning methods. We will also discuss how to achieve reliable machine learning by exploiting causal knowledge from observational data.

Causal Inference Applications

In this part, we will discuss how the presented techniques can be applied to real-world application scenarios. For example, in the field of digital marketing or online advertising, how to maximize the possible clicks by applying treatment effect estimation. Another example is to incorporate the idea of causal inference into the recommendation system.

Meanwhile, we will also discuss how real-world applications motivate new research topics and enlarge the scope of causal inference. For example, most of the existing methods deal with categorical/numerical data types, but textual covariates are commonly observed in real-world applications such as advertising. Motivated by such great demand, recent work starts the estimation of treatment effect with textual covariates.

Conclusions and Future Perspectives

We will conclude this tutorial and point out a list of open problems and future research directions.