RESEARCH STATEMENT
Ruocheng Guo (rguo12@asu.edu)
Arizona State University
http://www.public.asu.edu/~rguo12/

Causal inference - that investigates the causal relationships among variables, and machine learning - that focuses on learning statistical associations among variables from observational data, are two of the most important components of data-driven research across various areas including data science, economics and statistics etc. To bridge the gap between causal inference and machine learning toward causal machine learning, two natural questions can asked: (RQ1) how can we leverage advanced machine learning techniques to address challenges in causal inference problems? (RQ2) How can causal knowledge be utilized to improve solutions to machine learning problems? The big data era grants researchers with convenient access to massive observational data. On one hand, massive observational data provides the opportunities to investigate various influential causal machine learning problems without performing randomized controlled trials (RCTs) that can be expensive, time consuming and even unethical. Regarding RQ1, we can guide decision making by (RQ1.1) estimating the causal effect of a treatment (e.g., lockdown) on an important outcome (e.g., the number of COVID-19 cases) and (RQ1.2) evaluating and optimizing the utility of treatment assignment policies. For RQ2, we aim to design causal inductive bias for improving machine learning models trained with observational data through embedding the limited but important prior causal knowledge. We can design different causal inductive bias for different goals. In (RQ2.1) out-of-distribution generalization, properly designed causal inductive bias can improve models’ generalizability in unseen domains. In (RQ2.2) unbiased interactive machine learning, we aim to optimize models’ performance in online A/B tests while training them using offline data with estimated online metrics. On the other hand, massive observational data can be high-dimensional, multi-modal and heterogeneous. Such characteristics lead to limited prior knowledge of the underlying causal model of a given dataset. This fact poses challenges in the following facets. (C1) Confounding bias is the bias in estimated causal effects and policy utilities due to the existence of confounders, i.e., the variables causally influence both treatment and outcome. (C2) Selection bias refers to the cases where a subset of instances has less chance to be sampled into the observed population due to preferential selection. (C3) Spurious correlations are the statistical associations among variables which can be unstable across different subpopulations (domains or environments) in heterogeneous data.

Research Theme: My research aims to bridge the gap of causal inference and machine learning by developing a suite of novel causal machine learning algorithms to understand, characterize, and gain actionable insights from massive observational and experimental data, to benefit high-impact real-world applications.

1 Research Contributions

My research work boils down to developing principled algorithms to tackle the aforementioned challenges (C1, C2, and C3) for solving both causal inference and machine learning problems with massive observational data. Specifically, we developed algorithms in two directions: (i) machine learning for causal inference - by leveraging advanced machine learning techniques for mitigating confounding bias in causal inference with networked observational data; and (ii) causality-aware machine learning - by designing causality-aware machine learning algorithms to alleviate selection bias and spurious correlations. Consequently, causality-aware machine learning algorithms can be robust against selection bias in training data and are expected to generalize to out-of-distribution test data. An overview of my research is summarized in Figure 1.

1.1 Machine Learning for Causal Inference with Networked Observational Data

Network information showing the topology of instances is ubiquitous in observational data such as social networks, graph of products in an e-commerce website and a spatial network of geometric objects. We are also interested in the effect of an intervention to outcomes in such data. We know that observational data poses the challenge of confounding bias in causal inference problems. If the influence of confounders is not properly controlled, the biased estimates of causal effects and policy utilities can lead to suboptimal decisions. For example, a local government may end up with a financial aid assignment policy that only achieves a suboptimal effect on the regional economical system. To handle confounding bias, it is imperative to leverage advanced machine learning techniques to learn high-quality latent confounders from networked observational data. In particular, we exploit the connection patterns among individuals from the auxiliary network information of observational data to learn latent variables that compensate for the hidden confounders. Conventional causal inference methods cannot handle hidden confounders as they overwhelmingly rely on
the untenable assumption of ignorability (a.k.a. unconfoundedness) that there does not exist hidden confounders. In this research, we developed a series of machine learning based causal inference algorithms to tackle this fundamental research issue in causal inference with networked observational data. Considering that social networks of individuals often embed patterns of some important social and behavioral hidden confounders (e.g., socio-economic status), we proposed a graph convolutional network based causal inference framework Network Deconfounder [1]. Specifically, Network Deconfounder learns latent variables to compensate for the hidden confounders using the network information among instances. To ensure that the learned latent confounders mitigate the confounding bias, we propose to balance the distributions of latent confounders on the treatment group level. The effectiveness of latent confounders learned by Network Deconfounder is demonstrated by the superior performance in the task of causal effect estimation with networked observational data. Later on, we extended the idea of learning latent confounders for mitigating confounding bias in counterfactual evaluation of treatment assignment policies [2]. In this task, given networked observational data, we aim to estimate the expected utility of a policy which assigns treatments to instances. Our proposed method CONE is motivated by the idea of mitigating confounding bias with latent confounders and the doubly robust estimator. Our empirical results lead to two implications. First, we can extract latent confounders by maximizing the mutual information between the latent variables which are predictive of treatments and those that are predictive of outcomes. Second, we can improve the performance of doubly robust estimator due to the use of high-quality latent confounders in both the effect estimation model and the propensity model. Meanwhile, conventional methods assumes that all confounders are readily available in the observational data, while the assumption is not always tenable as we are often confronted with unknown causal model and hidden confounders. In [3], we made one of the first attempts of adversarial learning for extracting latent confounders from networked observational data. We develop a framework called IGNITE by designing a minimax game that can benefit from two types of widely adopted latent confounder learners: those learn latent confounders from propensity score modeling and those learn them by representation balancing. We empirically show that the proposed IGNITE framework is consistently and significantly outperforming existing methods in causal effect estimation with networked observational data.

**Impact/Results.** Our survey paper on learning causality with observational data [4] has appeared in ACM Computing Surveys in 2020. Since its publication, it attracts more than 40 citations. An open source causality algorithm repository - awesome-causality-algorithms (https://github.com/rguo12/awesome-causality-algorithms/) has been developed which consists of around 80 popular causal learning algorithms for different types of data. It has received been featured on several blogs, and has attracted more than 900 stars on Github. We are one of the first to leverage graph neural networks [1] and variational autoencoders [5] for causal inference in networked observational data.

### 1.2 Causality-aware Machine Learning

Making machine learning models causality-aware is crucial in a variety of important research problems including (1) unbiased offline interactive machine learning and (2) out-of-distribution generalization.

**Unbiased offline interactive machine learning.** The challenge of this problem is that when interactive machine learning models (e.g., search ranking and recommender systems) is trained with user generated labels (e.g., clicks and purchases) we can only train the model in an offline fashion with historical batch data but the goal is optimize the models’ performance in online A/B tests. To achieve this goal, inverse propensity scoring (IPS) based methods have been proposed. The propensity score of a user generated label is defined as the probability of the corresponding item (e.g., a product in an e-commerce search result page) to be examined by the user who generated the label. The main challenge in IPS based approaches is that it is often challenging to correctly measure the propensity scores without performing expensive and time consuming randomized controlled trials. In [6], we find that in e-commerce search, products are displayed in two-dimensional grids, which can result in significantly different user examination from the traditional web search with one-dimensional list display. Therefore, we propose to design propensity models for e-commerce product search ranking by utilizing the prior knowledge of user examination patterns in two-dimensional display. Empirically, we show that search ranking algorithms trained with the proposed propensity models lead to significantly better search ranking performance evaluated with real-world e-commerce search logs where there exists a distribution shift between the training and test data to mimic an online A/B test.

**Out-of-distribution generalization.** In this problem, the goal is to learn deep learning models that can generalize to unseen domains. We can consider this problem as a special case of selection bias. Leveraging the robustness of causal relationships across domains, we extend the invariant risk minimization (IRM) principle to learn causal feature representations for out-of-distribution generalization. In particular, we find that when strong spuriousness exists among the domain variable, the spurious features and the label, the original IRM principle can be satisfied by undesired solutions, which results in models that pick up spurious correlations. Therefore, we propose a simple but effective fix, namely IRM-CDM, which combines the conditional distribution matching constraint with IRM. IRM-CDM can exclude the undesired solutions and ensure the deep learning models learn causal features. We designed a new dataset to show that IRM can be satisfied by spurious models when the correlations among domain labels, spurious features and target labels are strong. Our empirical results show IRM-CDM significantly improve the out-of-distribution generalization performance of deep learning models under the aforementioned strong spuriousness.

**Impact/Results.** We are one of the first to study unbiased offline interactive machine learning in two-dimensional display [6], and the results lead to a full paper at the prestigious SIGKDD Conference.
2 Future Research Directions

My current research on causal machine learning poses a wealth of fascinating but challenging research questions that I plan to address in the short-term (first 3-5 years). My long-term goal is to create human-level artificial intelligence by bridging the gap between causality and machine learning. My Ph.D. research is funded by NSF, ARL, and ONR. In the future, I will actively apply for research grants from NSF, ARL, ONR and other funding agencies. I will also seek collaborations with causal machine learning industries (e.g., Google X, Microsoft and Etsy) and researchers from relevant fields (e.g., geoinformatics, economics, experimental design and computational social science). This section outlines some future research directions that I am excited to pursue.

2.1 Short-Term Plan

Causal inference on bipartite graph with strong interference. Causal inference on graphs with strong interference comes with highly influential applications such as online advertisement and e-commerce. The main challenge is causal identification under interference. There are two types of existing methods to deploy randomized experiments for causal effect estimation. Clustering randomization methods perform clustering and then randomly assign treatments based on the clustering results. These methods overwhelmingly depend on clustering results. Recently, bipartite graphs are adopted to formulate this problem in a different way. In a bipartite graph, treatments can only be assigned to one set of instances (treatment nodes) but outcomes of interest are measured on a different set of instances (outcome nodes). In other words, there is strong interference between the two sets of nodes. For example, in a two-sided online marketplace, a platform (e.g., Amazon and Etsy etc.) can only apply treatments (e.g., discounts) on the seller side but we are interested in the buyers’ purchases decisions. In this setting, even with randomized experiments, we cannot guarantee the exposure of outcome nodes to the treatment is random since the underlying bipartite graph connecting treatment nodes and outcome nodes is not randomly generated. This fact makes techniques developed for observational data become valuable in this setting. Existing work in this area extends classic causal inference methods to bipartite graph experiments. However, important problems remain open. Can we estimate causal effects on the treatment nodes without performing additional experiments? Can we generalize the estimates to unseen subgraphs? The first problem can be solved by assigning treatments s.t. the causal effects on the two types of nodes can be identified. To achieve this, we attempt to formulate a unified view of clustering randomization and bipartite graph experiments. In the second problem, we propose to extend causal transportability to bipartite graphs with strong interference.

Causal machine learning in spatial time series. Spatial time series data has important applications in highly influential areas such as regional economics, computational agriculture and climate change. For example, accurate forecasting of climate change events such as drought and flood can save lives and reduce economical and environmental costs. There are at least two types of problems we consider: (1) forecasting and (2) classification. Existing work shows that causality-aware models can improve the accuracy of models by the robustness of causal relationships across time. In spatial time series, the main challenges include (a) the heterogeneity of the underlying causal models across space and (b) the sparsity of data in certain regions. To tackle the first challenge, I propose to design spatially adaptive causal learning models that can recognize spatial heterogeneity as an ensemble of global and local models. To tackle the second challenge, we propose causal meta learning paradigm, which aims to learn invariant causal relationships with the data efficiency of meta learning. It allows us to adapt a model pre-trained in a data rich region to the data sparse regions using few data points. Moreover, in such interdisciplinary research, it is also very important to deliver interpretable results to domain experts to get useful feedback and insights. Therefore, we propose to generate counterfactual examples for spatial time series classification and forecasting models to analyze their feature importance and sensitivity [7]. These research problems involve interdisciplinary research across causal machine learning, economics, and geoinformatics, which can lead to interdisciplinary work of high impact.

Weakly supervised causal machine learning. Causal machine learning problems such as (1) out-of-distribution (OOD) generalization and (2) algorithmic fairness are important steps toward System 2 Machine Learning. We solve these problems with novel causal inductive bias (e.g., Invariant Risk Minimization). The existing work often only considers clean data from a single modality (e.g., images or text). I propose to develop algorithms applicable to noisy data using two types of structures: (a) the causal graph that describes the noise generation process and (b) the relational structure showing connection patterns among instances. Relational structures that embed dependencies among instances are universal such as social networks and state grids. They can be used to infer latent variables as alternatives for hidden variables in the causal graph. At the same time, robustness of causal relationships can help us correct the noisy observations. When we investigate OOD generalization and fairness with weak supervision, with structural information, we can relax the assumption that domain labels are accurately measured. Instead they can be inferred from the relational structures. For example, in social media, to learn a machine learning model that generalizes to different communities in a social network, without observed community labels, we can infer which them using the patterns extracted from social relationships. With research in this direction, I aim to develop fair and OOD generalizable models with weak supervision data, which extends the range of applicable causal machine learning.

2.2 Long-Term Plan

Principled design of causal machine learning models. One of the ultimate goals of causal machine learning is to improve the machine learning models towards human-level prediction accuracy under complex environment (e.g., non-i.i.d. data). It is surprising that there still does not exist a systematic answer to the question: when does causality matter in machine learning problems? Given a specific condition, What is the principle solution to design causality-aware machine learning models? Recent work on causal machine learning makes nice touches. For example, in recent work of causal machine learning, people often interpret causal relationships as invariant relationships. The conflict between this interpretation and Pearl’s structural causal models (SCMs) should be noticed.
In principle, with SCMs, we should be able to write down any causal concept with a non-parametric SCM. However, invariance is defined as a parametric condition. Given a SCM, we must be able to explain invariance by non-parametric constraints such as conditional independence. Moreover, researchers debate whether augmenting data with counterfactual examples helps machine learning models generalize to out-of-distribution data. The current debate is based on empirical results without a principled analysis with the SCMs of the datasets. To avoid such confusion in the machine learning community, I propose to build a systematic view of (1) how these concepts (e.g., invariance and counterfactual examples) can be formally written as conditions of SCMs and (2) derive design principles of causal machine learning models based on the implications from SCMs. This research direction provides us solid theoretical support for any causal machine learning model developed based on new concepts to solve real-world problems.

**Coupling causal machine learning with domain knowledge.** In many scientific and engineering areas, it is difficult to collect big data, but there exist physics-based models that embed domain knowledge. In fact, a physics based model is a structural causal model, it does not only describe the causal relationships among variables but also specify functions for each causal relationship. However, there are two major drawbacks of physics based models: (1) the expensive computational cost to sample data instances and (2) the mismatch between the model and the real-world data. Recently, machine learning researchers study how to connect machine learning to physics-based models. For example, recurrent neural networks are used to model the conditional intensity function of point process [8]. To overcome the issue (1), black box generative models are adopted to mimic the behavior of physics-based models. For issue (2), existing work either (a) only uses a small part of the physics system (e.g., a well defined physics law) as constraints of the machine learning model or (b) generates samples from a physics-based system to pre-train machine learning models without considering the mismatch problem. These approaches do not consider causality implied by physics based models and can result in machine learning models with spurious correlations. For issue (1), we can design causal inductive bias as part of the deep generative model to avoid spurious correlations. For example, if we know $X$ is the cause of $Y$, then to generate samples from $P(X, Y)$, we can follow the independent mechanism to design a deep latent variable model that first samples $X$ from $P(X)$ and then samples $Y$ from $P(Y|X)$. For issue (2), we propose to find a set of diverse simulation settings. This results in the multiple data distributions sharing the same robust causal relationships, which can be learned by machine learning models. We can use the generative model in (1) to reduce the computational cost in (2).

**Promoting interdisciplinary research.** My rich interdisciplinary collaboration experience with researchers from both academia and industry expands the scope of my research from data science and machine learning to other highly influential areas such as economics, geoinformatics, and social science. My experience shows the great potential of causal machine learning in interdisciplinary research problems. Novel research ideas are inspired by such collaborations, which can lead to the broad impact of my research. During my internship at Google X, my teammates are researchers with various backgrounds from geoinformatics, economics to business strategy. Our team made significant progress on an early stage confidential project. Moreover, with Dr. Deborah Hall from the School of Social and Behavioral Sciences of Arizona State University, we developed principled machine learning models for cyberbullying detection using hierarchical text data collected from social media [9]. I am now working with an economist who leads the online experimentation platform of a company, trying to use machine learning for experimental design in real-world business scenarios. Computational health is another interdisciplinary direction in my plan due to its strong connection with causal inference.

**Selected References**


