

# Welcome to MS&E 328/CS 328 – Foundations of Causal Machine Learning

Instructor: [Vasilis Syrgkanis](#), Assistant Professor, MS&E and CS (courtesy)

Units: 3

Spring Quarter 2023

Tue, Thu 3:00-4:20PM Littlefield Center, Rm 107

## Description and Format:

This is an advanced PhD course on modern theoretical topics at the intersection of causal inference, econometric theory and statistical learning theory. The course will consist of lectures, and student-led presentations. Each week will roughly consist of one lecture and two or three student presentations of research papers related to the theme of the week. For both parts, all students will need to read the corresponding papers and write short reviews of them. The course will also consist of a quarter long project which will involve a detailed literature review on a topic and either a replication of the methodology and the experimental results of a recent research paper or new methodological or theoretical developments on the topic of the project. Candidate research project ideas and areas will be provided; students are also free to propose other projects. The projects can be individual or in teams of up to three students.

**Prerequisites:** Good understanding on theoretical foundations of probability and statistics. Optimally, some familiarity with causal inference.

## Office Hours: (Starting Week 2)

	Time	Location
Vasilis Syrgkanis	Wednesday 2.30-3.30pm	Huang 252

## Grading

- 35% class presentation(s) / paper summaries
- 15% scribing
- 50% project (20% literature review, 30% final report)

## Course webpages

- Discussion and homework material: [canvas](#)
- Submissions:

## Late day policy

To accommodate unforeseen challenges that may arise during the quarter, you have three late days for the problem sets. Each late day allows you to turn in an assignment up to 24 hours late. (Any fraction of a late day counts as one late day.) You may use multiple late days on the same problem set. Work submitted beyond the allowed late days will not receive credit.

Please note that we have provided the late day policy to help provide flexibility to you in managing your course load during the quarter. If circumstances arise that require further accommodations, we encourage you to contact your academic advisor as well as the Office of Accessible Education (see below) to help make appropriate arrangements. Out of fairness to all students, in the absence of an OAE Academic Accommodation Letter, we will generally be unable to provide accommodations beyond the late day policy above.

## Access and accommodations

Stanford and our class are committed to providing equal educational opportunities for disabled students. Disabled students are a valued and essential part of the Stanford community. We welcome you to MS&E 228.

If you experience disability, please register with the Office of Accessible Education (OAE). Professional staff will evaluate your needs, support appropriate and reasonable accommodations, and prepare an Academic Accommodation Letter for faculty (not only for your teaching staff in MS&E 226, but also your other courses). To get started, or to re-initiate services, please visit

[oae.stanford.edu](http://oae.stanford.edu).

If you already have an Academic Accommodation Letter from OAE, we invite you to share your letter with us. Academic Accommodation Letters should be shared at the earliest possible opportunity so we may partner with you and OAE to identify any barriers to access and inclusion that might be encountered in your experience of this course.

### Honor code

Your work on problem sets and the project is governed by the [Stanford Honor Code](#). Any violations of the Honor Code will be referred to the Office of Community Standards for adjudication.

### Submitting on Gradescope

As noted above, problem sets and project submissions are submitted and graded through Gradescope. To ensure this process is smooth, there are a few things to keep in mind:

- You are required to tag your answers correctly. The graders will ignore any part of your solution that is not tagged. Note that this means you also have to correctly tag your code. Allow enough time prior to submission to ensure you are able to tag correctly.
- In order to grade code you submitted, we need to be able to copy your code. Make sure this is possible (e.g., do not upload screenshot images of your code). We will deduct points if we cannot check your code.
- If you believe we have made a mistake grading your work, you should submit a regrade request through Gradescope. This sends your request directly to the grader on that particular question. You must submit your regrade request within 14 days of the grades of that particular problem set or project part being published.

### Course communications: Ed Discussion

As noted above, we will use Ed Discussion to manage course announcements and a discussion forum. Ed Discussion will be available through Canvas.

Please use Ed Discussion for all course-related communication with us. We will aim to respond to questions in a 24-48 hour period, except of course for those of an urgent nature (e.g., typos on problem sets or lecture notes, clarifying course logistics, etc.). We encourage students to respond to each other, particularly during this “waiting period” before course staff answers! Among other things this waiting period means you should not wait until the last day before a problem set is due to message us; we may not be able to respond in time.

### Office hours

You are encouraged to attend office hours to ask questions of a technical nature. Office hours will be held in-person. Our goal is to foster an inclusive environment for additional learning support during office hours.

If you are having difficulty, we urge you to seek help from the TAs or the instructor as soon as possible. The material builds on itself and a solid understanding of the foundations is necessary for the rest of the course. Remember, the teaching team is here to serve as a resource for any questions you may have.

### Inclusion

It is our intent that students from a diverse set of backgrounds and with a wide range of perspectives be well served by this course; that students’ learning needs are being addressed both in and out of class; and that the diversity that students bring to this class becomes a resource, strength, and benefit to all of us. We try to present materials and lead discussions that are respectful of the diversity in our student population. Your suggestions are encouraged and appreciated; please let us know of ways to improve the effectiveness of the course for you personally or for other students or student groups.

Unfortunately, incidents of bias, discrimination, or intolerance do occur, whether intentional or unintentional. They can contribute to creating an unwelcoming environment for individuals and groups in the classroom. Please speak out if such an event occurs and we will do our best to handle it accordingly. You can reach out to the teaching team directly or bring the concerns to the University’s [Diversity and Access Office](#) and Acts of Intolerance to [Student Affairs](#).

## Suggested textbooks

Course lecture notes will be posted online in the form of powerpoint presentations. In addition accompanying chapters of a course textbook will be released on Canvas. Any feedback on slides and the course textbook will be greatly appreciated as we iterate over the material. These are the primary material for the course. In addition, you may find it helpful to have the

following textbooks on hand. I'm not requiring them, because they are available online, and we will not be linearly working our way through any of them.

- Hernan MA, Robins JM, [Causal Inference: What If](#)
- Scott Cunningham, [Causal Inference: The Mixtape](#)
- Kiciman, Sharma, [Causal Reasoning: Fundamentals and Machine Learning Applications](#)
- Shai Shalev-Shwartz and Shai Ben-David, [Understanding Machine Learning: From Theory to Algorithms](#)
- Wainwright, [High Dimensional Statistics: A Non-Asymptotic Viewpoint](#)

## Other references

Here is an assortment of other books that you may find useful to consult but which are not available online.

- Judea Pearl, [Causality](#)
- Pearl, Glymour, Jewell, [Causal Inference in Statistics](#)
- Pearl, Mackenzie, [The Book of WHY: the new science of cause and effect](#)
- Peters, Janzing, Schölkopf, [Elements of Causal Inference](#)
- Matt Taddy, [Business Data Science](#)
- Kohavi, Tang, and Xu, [Trustworthy Online Controlled Experiments](#)
- Imbens and Rubin, [Causal Inference for Statistics, Social, and Biomedical Sciences](#)
- Angrist and Pischke, [Mostly Harmless Econometrics](#)

## Course Plan

**Lecture 1:** Potential Outcomes and DAGs.

Causal identification via the potential outcomes framework, the structural equation framework and its DAG representation. Single world intervention graphs (SWIGs) and identification by adjustment. Proof of D-separation implies conditional independence.

**Lecture 2:** Automated Identification in DAGs: the ID Algorithm

The ID algorithm for general identification and its proof of sufficiency and necessity.

**Lecture 3:** Identification with Unobserved Confounding beyond DAGs

Causal identification in the presence of unobserved confounding beyond DAGs. (Instrumental variables, DiD, RDD, Synthetic Controls, Proximal inference).

**Lecture 4:** Semi-Parametric Inference and Neyman Orthogonality

Estimation via moment conditions. Neyman orthogonality and Debiased Machine Learning. Proof of asymptotic linearity for Neyman orthogonal moments with sample splitting and without sample splitting. Automatic Debiased Machine Learning and proof. Proof of the Lasso rate. Discuss the multiplier bootstrap for joint inference.

**Lecture 5:** Identification and Inference in Dynamic Regimes

Identification and inference of dynamic counterfactual policies (non-parametric and semi-parametric). Estimation of optimal dynamic regimes and g-estimation. Identification proof. Auto-DML and proof for dynamic regime

**Lecture 6:** Orthogonal Statistical Learning Theory for Heterogeneous Effects

Orthogonal statistical learning theory. Localized rademacher complexities and generalization bounds.

**Lecture 7:** Non-Parametric Inference for Heterogeneous Effects

Non-parametric confidence intervals, random forests and nearest neighbors. Proof of asymptotic linearity for kernel based moment estimators. Proof of bias for k-NN and (maybe proof of bias for Trees). Proof of confidence intervals with nuisance parameters and local orthogonality.

**Lecture 8: Non-Parametric Learning and Conditional Moment Restrictions**

Adversarial estimators for conditional moment restrictions. Statistical learning theory for adversarial estimators. Confidence intervals on functionals of endogenous regression functions. Proof of the rate for adversarial estimators based on the localized complexities. Proof of the auto-debiased approach for functionals of endogenous regressions.

**Lecture 9: Sensitivity Analysis and Causal ML**

Omitted variable bias in semi-parametric and non-parametric models. Inference on bias bounds

**Lecture 10: Representation Learning and Causal Inference**

Linear and Non-Linear Independent Component Analysis. Impossibilities and possibilities in non-linear causal representation learning.

**Lecture 11: Causal Discovery and Post-Discovery Inference**

Linear ICA and discovery; LinGAM (proof of identification). Causal discovery with unobserved confounding (FCI). Conditional independence testing.