Carleton University

ECON 4709/ECON 5880 W: Economic Data Science - Applications

Winter 2024

Professor Thomas Russell

General Information

The Instructor

Instructor: Thomas Russell

Email: thomas.russell3@carleton.ca

Office Location: Loeb Building A806

Office Hours: By appointment

The Course

Course Delivery: In person

Lectures: Monday 2:35 pm - 5:25 pm,View your timetable on Carleton Central to see the location.

TA: TBA

Brightspace Course Webpage: https://brightspace.carleton.ca/d2l/home/224922

Course Description: Application of data science and machine learning methods to real-world economic problems. Students will apply their data science knowledge in hands-on projects to answer topical research questions. This course has a strong practical focus.

Learning Objectives: The goal of this course is to demonstrate how machine learning algorithms can be used in economic applications, going beyond simple prediction tasks. The course will introduce students to a growing theoretical and applied literature in economics that uses machine learning tools to learn about causal or structural parameters, and to make optimal policy choices. Topics will include double/debiased machine

Dates	Event
January 8	Winter Term Begins
February 5	Assignment I Posted
February 19	Assignment I Due at 11:59 pm EST
February 19 - 23	Winter Break, no classes.
March 4	Empirical Project Proposal Presentations
March 18	Assignment II Posted
April 1	Assignment II Due at 11:59 pm EST
April 8	Empirical Project Final Presentations
	Winter Term Ends
	Empirical Project Written Submission Due at 11:59 pm EST

Table 1: Important dates for ECON 4709/5880.

learning, counterfactual proxies and synthetic control, learning optimal treatment assignment and optimal taxation and insurance policies, among others. The course will highlight the contributions of economists to the machine learning literature, and will emphasize applications.

There will be no written examinations in this course. Instead, students will have to complete two assignments as well as a major empirical project. For the empirical project, students must establish their own research questions, and must demonstrate mastery of the concepts taught in the course. Students will also need to present their project proposal and final findings to the class.

Course Preclusions: None.

Course Prerequisites: For ECON 4709, at least a C+ in ECON 4708. For ECON 5880 W, at least a C+ in ECON 5880 F.

Course Calendar

Table 1 displays a list of important dates. The evaluation dates below are subject to change.
The tentative course outline is displayed in Table 2. Depending on the pace of the lectures, this course outline may be subject to modification.

Lectures

Attendance of in-person lectures is considered mandatory, and class participation will be graded. Lecture recordings will not be available.

Topic #	Lecture Dates	Topic
1	January 8, 2024 January 15, 2024	Introduction: Prediction versus Explanation
2	January 22, 2024 January 29, 2024	Double/Debiased Machine Learning
3	February 5, 2024 February 12, 2024	Counterfactual Proxies and Synthetic Control Methods
—	February 19 - 23, 2024	Winter Break: No Classes
4	February 26, 2024 March 4, 2024	Welfare Maximization and Learning Optimal Policies
5	March 11, 2024 March 18, 2024	Machine Learning for Optimal Taxation and Insurance
6	March 25, 2024 April 1, 2024	Recent Research in Econometrics and ML
—	April 8, 2024	Empirical Project Presentations, Last Day of Class

Table 2: A tentative course outline for ECON 4709/ECON 5880.

Programming Requirements

Students will be required to use R for the assignments. Students will be required to use LaTeX for their empirical project.

Textbooks and Reading Materials

There is no required textbook for this course. Instead the lecture material will be drawn from a variety of different sources. Required readings below are marked with an asterisk *.

Topic 1: Introduction, Prediction versus Explanation

* Athey, S. (2017). Beyond prediction: Using big data for policy problems. Science, 355(6324), 483–485. https://doi.org/10.1126/science.aal4321

* Athey, S. (2019). The Impact of Machine Learning on Economics. In The economics of artificial intelligence (pp. 507-552). University of Chicago Press. nber.org/system/files/chapters/c14009/c14009.pdf

* Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical science, 16(3), 199-231.

- Kleinberg, J., Ludwig, J., Mullainathan, S., & Obermeyer, Z. (2015). Prediction policy problems. American Economic Review, 105(5), 491-95.

* Mullainathan, S., & Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. Journal of Economic Perspectives, 31(2), 87–106. https://doi.org/10.1257/jep.31.2.87

* Shmueli, G. (2010). To Explain or to Predict? Statistical Science, 25(3), 289–310. https://doi.org/10.1214/10-STS330

* Storm, H., Baylis, K., & Heckelei, T. (2020). Machine learning in agricultural and applied economics. European Review of Agricultural Economics, 47(3), 849-892.

- Varian, H. R. (2014). Big Data: New Tricks for Econometrics. Journal of Economic Perspectives, 28(2), 3–28. https://doi.org/10.1257/jep.28.2.3

Topic 2: Double/Debiased Machine Learning

- Belloni, A., & Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. Bernoulli, 19(2), 521-547.

- Belloni, A., Chernozhukov, V., & Hansen, C. (2014). High-dimensional methods and inference on structural and treatment effects. Journal of Economic Perspectives, 28(2), 29-50.

- Chernozhukov, V., Hansen, C., & Spindler, M. (2015). Valid post-selection and post-regularization inference: An elementary, general approach. Annu. Rev. Econ., 7(1), 649-688.

* Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters.

Topic 3: Counterfactual Proxies and Synthetic Control Methods

- Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature, 59(2), 391-425.

* Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2021). Matrix completion methods for causal panel data models. Journal of the American Statistical Association, 116(536), 1716-1730.

* Chernozhukov, V., Wüthrich, K., & Zhu, Y. (2021). An exact and robust conformal inference method for counterfactual and synthetic controls. Journal of the American Statistical Association, 116(536), 1849-1864.

* Doudchenko, N., & Imbens, G. W. (2016). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis (No. w22791). National Bureau of Economic Research.

Topic 4: Welfare Maximization and Learning Optimal Policies

* Athey, S., & Wager, S. (2021). Policy learning with observational data. Econometrica, 89(1), 133-161.

* Kitagawa, T., & Tetenov, A. (2018). Who should be treated? empirical welfare maximization methods for treatment choice. Econometrica, 86(2), 591-616.

* Mbakop, E., & Tabord-Meehan, M. (2021). Model selection for treatment choice: Penalized welfare maximization. Econometrica, 89(2), 825-848.

Topic 5: Machine Learning for Optimal Taxation and Insurance

* Kasy, M. (2018). Optimal taxation and insurance using machine learning—Sufficient statistics and beyond. Journal of Public Economics, 167, 205-219.

Topic 6: Recent Research in Econometrics and Machine Learning

TBD.

Evaluation

- Class participation — 10%. Marks will be deducted for missing class, or for failing to regularly engage with class discussion.
- Assignment I — 15%. To be posted on the course website on February 5, 2024. Due on February 19, 2024, at 11:59 pm EST. To be submitted via the course website.
- Assignment II — 15%. To be posted on the course website on March 18, 2024. Due April 1, 2024 at 11:59 pm EST. To be submitted via the course website.
- Empirical Project Proposal Presentation—5%. Each student must present for 15 min. Presentations will take place on March 4th during class.
- Empirical Project Final Presentation—10%. Each student must present for 30 min. Presentations will take place on April 8th during class.
- Empirical Project Written Submission—45%. Due on April 8, 2024 at 11:59 pm EST. To be submitted via the course website.

Assignments

Students may work in groups of one or two individuals. Only one assignment per group needs to be submitted – list all groups members on your assignments. Assignments will require students to apply concepts learned in

class to a mix of theoretical and applied questions. All assignments will be submitted via the course website. Students will submit both a written portion of the assignment (containing any mathematical derivations or explanations required to answer the assignment questions), as well as meticulously labelled R code used to answer any of the programming-based questions. Assignment marks may be deducted if any supporting code is not clearly labelled and/or explained. **Late assignments will be penalized 1 mark for each minute they are late (e.g. 5 minutes late means a 5 mark deduction from your total assignment score).**

The use of generative artifical intelligence (AI) tools (e.g. ChatGPT) is prohibited. Any use of generative AI tools to produce assessed content is considered a violation of academic integrity standards as per the statement on Plagiarism. Re-use of the same piece of work previously submitted for credit in another course is also prohibited.

Students registered in ECON 5880 will be required to answer additional assignment questions.

Satisfactory Performance Criteria

Students must fulfill all of the course requirements in order to achieve a passing grade.

Additional Information

Email Communication

Communication outside of class, tutorials and office hours will be done through email. **For security purposes, please communicate only using your carleton.ca email address.** Please include the course name "ECON 4709" or "ECON 5880" in the subject line of your email. You can expect to receive a response to your email within 48 hours. However, communication outside of class hours should occur only in exceptional cases; in particular, email is not the appropriate medium to ask questions about course material. Students who request clarification on course material through email will be directed to ask their question during class, tutorial or office hours.

Course Standing

Standing in a course is determined by the course instructor subject to the approval of the Faculty Dean. This means that grades submitted by the instructor may be subject to revision. No grades are final until they have been approved by the Dean. Application to write a deferred final examination must be made at the Registrars Office.

Copyright of Course Materials

The lectures and course materials (including all slides, handouts, recorded lecture videos, exams, tutorial materials, and other similar materials) are protected by copyright. The course instructor is the exclusive owner of the copyright and intellectual property of all course materials. You may take notes and make copies of course materials for your own educational use. You may not reproduce or distribute lecture notes, videos, or other course materials publicly without the express written consent of the instructor.

Plagiarism, Resources and Mental Health, Academic Accommodations

You are responsible for reading and knowing the information about plagiarism, Carleton University resources, and academic accommodations found here.