Carleton University

ECON 4708/ECON 5880: Economic Data Science - Analytics

Fall 2023

Professor Thomas Russell

General Information

The Instructor

Instructor: Thomas Russell
Email: thomas.russell3@carleton.ca
Office Location: Loeb Building A806
Office Hours: By appointment

Course Delivery

Course Delivery: In person. Log into Carleton Central to view the location on your timetable.
Lectures: Thursday 11:35 am - 2:25 pm

TA: TBA

Brightspace Course Webpage: https://brightspace.carleton.ca/d2l/home/212230

Course Description: An introduction to methods of statistical and machine learning analytics for economic analysis. Tools relevant for both small and large data sets will be covered. Topics may include approaches to classification, dimension reduction strategies, and prediction models and tools.

Learning Outcomes: The goal of this course is to introduce students to some of the most popular machine learning algorithms, and to demonstrate "off-the-shelf" applications of these algorithms to economic problems. The course assumes that students have sufficient background in statistics, probability theory, and multivariate linear regression. The course will focus primarily on supervised machine learning algorithms,

and may cover linear models, LASSO and ridge regression, logistic regression, nonparametric regression, decisions trees and random forests, support vector machines, bayesian methods, and neural networks. The course will also introduce students to programming in R, and will expose students to a variety of datasets from economics that are amenable to analysis using machine learning methods. No prior programming experience is required.

Course Preclusions: None.

Course Prerequisites: For ECON 4708, ECON 2708 with a grade of C+ or higher; and ECON 4706 (or equivalent) with a grade of C+ or higher. For ECON 5880, by permission of the department.

Course Calendar

Table 1 displays a list of important dates. The evaluation dates below are subject to change.

Dates	Event
September 7	First Class
September 28	Assignment I Posted
October 19	Midterm Exam (during lecture)
	Assignment I Due
October 23 - 27	Fall Break, no classes.
November 16	Assignment II Posted
December 7	Assignment II Due at 11:59pm EST
	Last Class
December 10 - 22	Final Exam Period

Table 1: Important dates for ECON 4708.

The tentative course outline is displayed in Table 2. Depending on the pace of the lectures, this course outline may be subject to modification.

Programming Requirements

Students will be required to use R for the assignments. Programming in R will be taught during lecture. R is a free and open-source programming language and environment used for statistical computing. Next to Python, it is one of the most in-demand languages for data science. Most of the programming in class will be done using Jupyter notebooks, although students should also install RStudio. Instructions on how to install R, RStudio and Jupyter notebooks will be posted on the course website before the first lecture. These programs should be installed before the first class.

Topic #	Lecture Dates	Topic	Readings
1	September 7, 2023	Introduction to Data Science, Intro to R	**(ISLR2)**: Ch. 1 & 2
2	September 14, 2023 September 21, 2023	Linear Regression, Shrinkage, and Model Selection	**(ISLR2)**: Ch. 3, 5 & 6
3	September 28, 2023	Nonlinear Extensions	**(ISLR2)**: Ch. 7.
4	October 5, 2023 October 12, 2023	Logistic Regression, Decision Trees, and Random Forests	**(ISLR2)**: Ch. 4 & 8
—	October 19, 2023	Midterm Exam (In class)	
—	October 23 - 27, 2023	Fall Break: No Classes	
5	November 2, 2023 November 9, 2023	Support Vector Machines	**(ISLR2)**: Ch. 9
5	November 16, 2023 November 23, 2023	Bayesian Methods	**(PRML)** Ch. 3.3 **(GPML)** Ch. 2
6	November 30, 2023 December 7, 2023	Neural Networks and Deep Learning	**(ISLR2)**: Ch. 10

Table 2: A tentative course outline for ECON 4708.

Textbooks and Reading Materials

Although we will use a variety of references in this course, our main references (not required) will be:

- **(ISLR2)** James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning with Applications to R, *Second Edition*. Springer New York.

- **(PRML)** Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer Science and Business Media.

- **(GPML)** Williams, C. K., & Rasmussen, C. E. (2006). Gaussian processes for machine learning (Vol. 2, No. 3, p. 4). Cambridge, MA: MIT press.

The lecture material will also draw from the following sources:

- Abadie, A., & Kasy, M. (2019). Choosing Among Regularized Estimators in Empirical Economics: The Risk of Machine Learning. The Review of Economics and Statistics, 101(5), 743–762. https://doi.org/10.1162/rest_a_00812

- 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

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media.

- Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). Foundations of machine learning. MIT press.

- 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

- Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press.

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

- 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

Evaluation

- Assignment I — 20%. To be posted on the course website on September 28, 2023. Due on October 19, 2023, at 11:59 pm EST. To be submitted via the course website.

- Midterm — 25%. The midterm will take place in class on October 19, 2023.

- Assignment II — 20%. To be posted on the course website on November 9, 2023. Due November 30, 2023, at 11:59 pm EST. To be submitted via the course website.

- Final Exam— 35%. The final exam will be administered during the exam period, sometime between December 10 - 22, 2023. The final exam will be scheduled later in the term and more details will be released at that time.

Assignments

Students can 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 programming skills 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, without exception. For instance, an assignment that is submitted 5 minutes late will receive a 5 mark deduction from the 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 relative to students registered in ECON 4708.

Midterm Exam

The midterm will be administered in class on October 19, 2023, and will cover lecture material up to and including the lecture on October 12, 2023.

There will not be a make up exam for those who miss a midterm. Those unable to complete the midterm for a valid reason will have the weight of the midterm shifted to the final exam, resulting in a final exam worth 60%. Students absent from the midterm should email the instructor as soon as possible. However, documentation may be required. The instructor reserves the right to request a Self-Declaration form or PMC letter of accommodation depending on the length of incapacitation. The documentation provided by students for missed or incomplete midterms will be subject to verification by the instructor. Submitting a forged note is regarded as a serious academic offense. Students who miss the midterm exam and fail to provide sufficient documentation will receive a score of zero, without exception.

Students registered in ECON 5880 will be required to answer additional midterm exam questions relative to students registered in ECON 4708.

Final Exam

The final exam will be administered in-person and will be scheduled by the University during the final exam period from December 10 - 22, 2023. The final exam will cover lecture material up to and including the final lecture on December 7, 2023. The exam will be scheduled later in the term, and relevant details for the final

exam will be announced at that time.

Students absent from the final exam should email the instructor as soon as possible. However, documentation may be required. The instructor reserves the right to request a Self-Declaration form or PMC letter of accommodation depending on the length of incapacitation. The documentation provided by students for missed or incomplete final exams will be subject to verification by the instructor. Submitting a forged note is regarded as a serious academic offense. Students who miss the final exam and fail to provide sufficient documentation will receive a score of zero, without exception.

Students are not to make travel plans during the exam period as that is not a valid reason for missing a final exam.

Students registered in ECON 5880 will be required to answer additional final exam questions relative to students registered in ECON 4708.

Satisfactory Performance Criteria

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

Deferred Final Exam

Students who do not write the final examination because of illness or other circumstances beyond their control may apply to write a deferred final examination by contacting the Registrar's Office no later than three working days after the original final examination was scheduled. In the event that a student writes a deferred examination, the deferred examination will carry the same weight as the final examination in determining the course grade. Any deferred examination will not be identical to the original final examination.

E-proctoring

E-proctoring will not be used for evaluations in this course.

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 4708" 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.