

ECO3400H1S

Economic Applications of Machine Learning

University of Toronto

Department of Economics

Winter 2021

Instructor: Marlène KOFFI

Mail: marlene.koffi@utoronto.ca

Office Hours: Thursday 5pm-6pm (Online)

Teaching Assistant: TBD

Overview

This is a graduate-level course in Machine Learning, with a special focus on tools relevant for economists. It is intended for students interested in applied research and/or analysis of big and unstructured data. Machine learning methods can be used to answer various questions in economics and in a plethora of subfields in applied research. Notable applications range from economics of education, labor, macroeconomics, crime economics, finance, econometrics, microeconomics, development, public economics, political economy, health economics, etc. The objective of this course is, therefore, at two levels. The first is to familiarize students with the most used machine learning methods in economics. The second is to cover cutting-edge economic research articles using these methods.

Course delivery

Lectures are Thursdays 3:10 pm to 5:00 pm (Toronto Time) on zoom.

The zoom link will be posted on Quercus.

The course mode will be “online synchronous”. The course will therefore be delivered during a livestream. There will be no video recording.

All course materials will be posted on Quercus. I recommend you check it regularly.

This course is intended to be interactive and to be able to mimic an in-person session at best. Therefore, students are expected to attend all class sessions, actively participate in discussions and activities. They must have a usable microphone and are encouraged to activate their camera. In case we manage to do presentations, it is required for the presenters to activate their cameras.

Time zone

All times posted will be in local Toronto time. If you are in a different time zone, please make sure you know of Toronto time in relation to your time zone. Errors in calculations are not an acceptable reason to miss deadlines.

Evaluation

Task	Weight	Date
Class Participation	15%	Every Thursday Lecture
Assignment	20%	April 1
Research Project (Group Work)		
<ul style="list-style-type: none"> • Part 1: Submission of ideas • Part 2: Final submission 	<ul style="list-style-type: none"> • 15% • 20% 	<ul style="list-style-type: none"> • February 25 • April 12
Final Exam	30%	Final Assessment Period

Class Participation:

This is a very interactive class. I expect you to engage in discussion constantly. All types of questions are welcomed and encouraged.

Assignment:

The assignment will be a python coding exercise. You will receive a sequence of three assignments. However, only one randomly selected exercise will be marked. More information will be provided during class.

Assignments are individual. You can help each other, but each student will have to submit their own “copy” and their own “code” in conformity to the Student Academic Integrity Code (See below).

Research Project:

You need to find a relevant question and apply machine learning tools to answer that question.

The submission of the project will have two parts. The first part is about *submitting the idea*. The submission of the idea will take the form of a written report which will **clearly** present the following steps:

- Context or motivation
- Contribution
- The method you plan to use
- A structured and concise literature review highlighting the similarities and dissimilarities between your work and existing work
- The data you plan to use.

This should not be longer than 5 pages.

The second part of the project, which is therefore the final submission, should be written in the form of papers published in AER: Insights or AER: Papers and Proceedings (both journals publish short papers) for an example of structure, length, etc.

The project will be individual for doctoral students and will be done in groups of 3 for master's students.

Final Exam

The final will be governed by the University's rules for final assessment. We will discuss it in class closer to the end of the semester.

Important Note: You are expected to work on the final exam by yourself, proctored or not. You may not use unauthorized aids or communicate with others about the test.

Late penalty:

All the evaluations are due at the "Due Date" indicated in the table, before 2:00 pm.

Late assignments will receive a grade of zero unless you make a declaration of absence via ACORN and notify the instructor (adding the TA in copy) before the due date.

Late research projects will be downgrade by 10 percentage points per day late.

If you don't take the final exam in the specified windows, you will receive a grade of zero unless you make a declaration of absence via ACORN and notify the instructor (adding the TA in copy) before the due date.

Course Materials

Textbooks

The core lecture material is based on:

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, p. 18). New York: springer. (JWHT)

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media. (HTF)

Other books

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.

Sarkar, D. (2016). *Text Analytics with Python*.

Key reviews:

Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda* (pp. 507-547). University of Chicago Press.

Athey, S. and G. Imbens “Machine Learning Methods Economists Should Know About”, *Annual Review of Economics*, 2019, Vol 11, 685-725.

Einav, L., & Levin, J. (2014). The data revolution and economic analysis. *Innovation Policy and the Economy*, 14(1), 1-24.

Evans, J. A., & Aceves, P. (2016). Machine translation: Mining text for social theory. *Annual Review of Sociology*, 42, 21-50.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3), 535-74.

Goldfarb, A., Gans, J., & Agrawal, A. (2019). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political analysis*, 21(3), 267-297.

Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87-106.

Varian, H. “Big Data: New Tricks for Econometrics”, Journal of Economic Perspectives, 28, 3-28, 2014.

Start with Python

Guido, S., & Müller, A. (2016). Introduction to machine learning with python (Vol. 282). O'Reilly Media.

Sweigart, A. (2019). Automate the boring stuff with Python: practical programming for total beginners. No Starch Press.

A non-exhaustive list of articles (Reading List) will be added to Quercus in addition to the aforementioned books and articles. The list of readings may be updated throughout the course. Please check it regularly.

Software

This course will use Python software exclusively. Python is a free software. All students are required to install it on their machine. However, Stata may be used for general economics/econometrics applications.

Since this is a graduate-level course, we will not have a step-by-step Python learning session (not enough time). You may have an introductory Python session with the TA. One of the TA's primary roles will also be to guide you when you have questions relating to Python and coding. However, you will have to learn and experience the python environment mostly on your own. In class, I will generally only give functions relating to a particular method.

Class Schedule

Date	Week	Topic	Chapters
January 14	1	Introduction	Chapter 2 of JWHT et HTF
January 21	2	Classification	Chapter 4 of JWHT et HTF
January 28	3	Developing Self-Learning Algorithm *Cross-Validation * Bootstrap	Chapter 5 of JWHT and Chapter 7 of HTF
February 4	4	Model Selection and Regularization	Chapter 6 of JWHT and Chapter 3 of HTF

February 11	5	Non-Linear Model	Chapter 7 of JWHT and Chapter 9 of HTF
February 18	Reading Week		
February 25	6	Tree-Based Methods SVM (if time allowed)	Chapter 8 of JWHT and Chapters 9 and 15 of HTF
March 4	7	Principal Component Analysis and Clustering	Chapter 10 of JWHT and Chapter 14 of HTF
March 11	8	Natural Language Processing and Feature engineering Model	Sarkar, D. (2016) Chapters 3 and 4
March 18	9	Topic Modelling and Sentiment	Sarkar, D. (2016) Chapters 6 and 9
March 25	10	Neural Network	Chapter 11 of HTF
April 1	11	Causal Inference with Machine Learning	Athey (2018) Athey and Imbens (2019) Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2018). <i>Generic machine learning inference on heterogenous treatment effects in randomized experiments</i> (No. w24678). National Bureau of Economic Research.
April 8	12	Presentation of Project (If time)/ Discussion /Time for any necessary catch-up	

Email Policy

Please feel free to email me questions or comments pertaining to the course, with the following caveat:

The answer requires a one or two-line response. It is my experience that email is an inefficient way to discuss economics. Questions that require more than one or two-line answers are more appropriate for office hours.

Students are also required to send emails using their University of Toronto email addresses. The subject of the email should include the course number and the UTORid.

Academic Integrity

Students should note that copying, plagiarizing, or other forms of academic misconduct will not be tolerated. Any student caught engaging in such activities will be subject to academic discipline ranging from a mark of zero on the assignment, test or examination to dismissal from the university as outlined in the academic handbook. Any student abetting or otherwise assisting in such misconduct will also be subject to academic penalties.

Moreover, you are expected to know, review and apply all University requirements regarding academic integrity. Please read:

<https://www.artsci.utoronto.ca/current/academic-advising-and-support/student-academic-integrity>

<https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019>

University disclaimer concerning Turnitin:

“Normally, students will be required to submit their course essays to Turnitin.com for a review of textual similarity and detection of possible plagiarism. In doing so, students will allow their essays to be included as source documents in the Turnitin.com reference database, where they will be used solely for the purpose of detecting plagiarism. The terms that apply to the University's use of the Turnitin.com service are described on the Turnitin.com web site.”

Class Materials Policy:

Class materials are subject to the University's policy on intellectual property. It is strictly forbidden to copy, share, or distribute any class materials, except for the current academic use purpose, without the instructor's explicit permission.

Code of Conduct in an online environment

- The thing number one to recall is that we are in a learning environment. Mistakes, discussions, exchange of ideas, etc., are fine as long as there are made in total respect of the person and individuality.
- Please, mute yourself unless you need to talk for the class's benefit (ask a question, answer a question, etc.).
- To avoid unpleasant interruption, when you want to ask a question, please use the chat function or wait for the time allowed to do so. As you will notice, I frequently ask if there are any questions. When asked, you are welcomed to unmute yourself and ask any question you may have.
- Again, I make a point of honor to have a respectful environment during class. So please, respect your peers. Use proper and respectful language and refrain from any insults, threats or bad jokes.
- Finally, adhere to the same standards as you would in the classroom.

Academic Accommodations

The University is committed to accessibility. If a student requires accommodations for a disability, or has any accessibility concerns about the course, please contact Accessibility Services as soon as possible. Their website is <http://www.studentlife.utoronto.ca/as>.

Other important notice

The students are expected to comply with all University policies even if not expressly mentioned above.