University of Waterloo
Department of Economics
Course Outline
Winter 2019

Econ 422: Microeconometric Analysis:
An Introduction to Machine Learning for Economists

Statistical Learning	Lasso	Conjunctive Concepts	Bagging	Regularization	Training, Testing, Validation	Features	*Uber*
No Free Lunch Theorem	Machine Learning	Random Forest	Boosting	Five Tribes	AlphaZero	*YouTube*	Selection Bias
Unsupervised Learning	Overfitting	Neural Network	Support Vector Machines	Decision Trees	*LinkedIn*	Difference-in-Differences	Treatment and Control
Supervised Learning	Tank Problem	A/B Testing	Prediction	*Google*	Regression Discontinuity	Identifying Variation	Mean
Text Analysis	Artificial Intelligence	Recommender Systems	*Your life, your job*	Credibility Revolution	p-Hacking	Logit Model	Competing Economic Hypotheses
Cloud Computing	Neats and Scruffies	*Amazon*	Randomized Controlled Trials	Null and Alternative Hypothesis	Causal Effect	External Validity	Counterfactual
Parallel Processing	*Netflix*	Instrumental Variables	Omitted Variable Bias	Cluster-Robust Standard Errors	Endogeneity	Econometrics	Statistical and Economic Significance
Facebook	Regression	Natural Experiments	Size	Power	Monte Carlo	Bootstrap	Conditional Expectation

Course Introduction: This course is an introduction to machine learning for economists. Our study of machine learning algorithms will focus mainly on supervised learning methods but we will also cover some unsupervised learning approaches. We will also compare the aims and tools in machine learning vs. traditional econometric analysis. Additional topics include using text as data, utilizing cloud computing, online recommender systems, and the impact of machine learning on society. We also review the role of causal empirical analysis in economics and the five main econometric methods.

Course Objective: The main objective is to provide students an opportunity to become familiar with key machine learning concepts and methods and how economists are using machine learning in applications.

Instructor: Mikko Packalen, Associate Professor

Course Pre-requisites: Familiarity with principles of data analysis and causal inference (<u>one</u> of ECON 321, ECON 323, STAT 221, STAT 231, STAT 241).

Office Hours: <u>Usually</u> on Thursdays 1am-3pm in HH 205, but please check the course website for possible changes (learn.uwaterloo.ca); and by appointment (please email/talk after class to make one).

Email: packalen@uwaterloo.ca; subject line must state "ECON 422" and include your full name and student ID number and the email must come from your .uwaterloo account (this way I know that it's really you and that it's an important email).

1. Course Materials

The core course materials are: Lectures, Problem Sets and Quizzes, Scientific Journal Articles, Policy Papers and News Paper Articles, Textbooks, and Computer Software (R, or Stata and Python). These materials are <u>complementary to one another</u>; none of them is a substitute to another.

2. Lectures

Tuesdays and Thursdays, 4-5.20pm, RCH 204.

3. Tentative List of Topics

ML AND ECONOMETRICS
1. **The Credibility Revolution: What Economists Do with Data**
2. **Econometrics vs. Machine Learning: Parameter Estimation vs. Prediction**

ML ALGORITHMS I
3. **Model Evaluation** (ISLR 2.1, 2.2.1-2.2.2, 2.3,5.1.1-5.1.3,5.2,5.3.1-5.3.3)
4. **Regression Trees** (ISLR 8.1, 8.3.1-8.3.2)
5. **Penalized Regression** (ISLR 6.1-6.2, 6.5-6.6)
6. **Classification: KNN and LDA** (ISLR 2.2.3, 4.1-4.6)
7. **Random Forests, Bagging, Boosting** (ISLR 8.2, 8.33-8.34)

DATA AND COMPUTING TOOLS
8. **Text as Data**
9. **Parallel Processing and Cloud Computing**

ML ALGORITHMS II
10. **Neural Networks**
11. **Unsupervised Learning** (ISLR 10.3, 10.5.1-10.5.2)
12. **Classification: SVM** (ISLR 9.1-9.3)
13. **Recommender Systems**

ML AND SOCIETY
14. **Data Ownership and Privacy**
15. **Machine Learning, AI, and Economic Growth**

4. Problem Sets

Problem sets are distributed through learn.uwaterloo.ca. Tentative list of problem set topics:

Problem Set 1: Regression Trees
Problem Set 2. Penalized Regression
Problem Set 3. Classification
Problem Set 4. Random Forests, Bagging, Boosting
Problem Set 5. Text as Data

[More advanced topics are covered in class and in the term paper].

5. Term Paper

The term paper project is an opportunity for students to gain experience in empirical analysis. Ideally, this research experience and the resulting paper will be beneficial to students in job interviews. The term paper should be 7-12 pages long (main text only, double-spaced). The empirical analysis can be either original analysis or a replication analysis of an existing study. However, in both cases all writing must be student's own. Further details will be discussed in class.

6. Software

R and Python are **free**. Stata is available at a discounted rate for registered students. Links to be added.

7. Textbooks and Scientific Articles

We will use the following **free** textbook on machine learning:

[ISLR] James G., Witten, D., Hastie, T. and R. Tibshirani (2013) *An Introduction to Statistical Learning with Applications in R.* Springer. Available at: http://www-bcf.usc.edu/~gareth/ISL/

We will also use various **freely accessible** scientific journal articles and book chapters, including:

Varian, H. R. (2014) "New Tricks for Economists," *Journal of Economic Perspectives*, 28(2): 3-28. Available at: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.28.2.3

Mullainathan, S. and J. Spiess (2017) "Machine Learning: An Applied Econometric Approach," *Journal of Economic Perspectives*, 31(2): 87-106. Available at: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.31.2.87

Domingos, P. (2012) "A Few Useful Things to Know About Machine Learning," *Communications of the ACM*, 55(10): 78-87. Available at: https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf

Domingos, P. (2015) "The Machine-Learning Revolution," chapter 1 in *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World.* Basic Books. Available at: https://books.google.ca/bkshp?hl=en&tab=pp

Athey, S. (2018) "The Impact of Machine Learning on Economics", to appear in Agrawal, A. K., Gans, J. and A. Goldfarb (eds.) *The Economics of Artificial Intelligence.* University of Chicago Press. Available at: https://www.nber.org/chapters/c14009.pdf

Angrist, J. D. and J.-S. Pischke (2017) Undergraduate Econometrics Instruction: Through Our Glasses, Darkly. *Journal of Economic Perspectives*, 31(2): 125-44. Available at: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.31.2.125

Christopher, R. (2018) "Shackling the Identification Police," Manuscript. Available at (from campus only): https://www.nber.org/papers/w25320

During the course we will also review the five main methods of causal inference. Many textbooks include this material; the following textbooks are the recommended (but **optional**) material for studying these topics (copies will be placed on reserve at Dana Porter Library):

Angrist, J. D. and J.-S. Pischke (2015) *Mastering Metrics: The Path from Cause to Effect.* Princeton University Press.

Stock, J. H. and M.W. Watson (2011) *Introduction to Econometrics*. Pearson.

8. Assessments

Final grade components and weights:

In-Class Participation and Knowledge	40%
Problem Sets	25%
Term Paper	35%

Notes:
- The in-class participation and knowledge component is measured by presence in class (75%) and by extent and quality of contributions to class discussions and accuracy of quiz answers (25%).
- Term Paper is due on April 20th, no late submissions are accepted.

9. Course Description (from Course Calendar)

This course investigates advanced estimation and inference techniques for microeconomic data. Students will learn about error components models that are used for economic data that exhibits significant unobserved heterogeneity. The estimation of treatment effects using fixed effects and difference-in-differences methods will be covered, and design-based methods of causal inference such as matching and regression discontinuity may be covered as well. Extensions such as multilevel or hierarchical models, limited dependent variable models, duration models or selection models may also be included. Students also learn how to apply these methods using computer software, and will use it to analyze complex data from household and firm-level surveys in assignments.

10. Additional Information

1. Missing a Problem Set Deadline or Quiz Due to Illness During Term
   - If a student misses a problem set deadline or quiz due to illness and has valid documentation (approval required), the weight of the missed problem set or quiz will be shifted to other problem sets. Without valid and timely submitted documentation, the student will receive zero for the missed problem set/quiz.

2. Fee-Arranged Issues
   Students are responsible for administrative matters concerning their course registration including fee arrangements. No make-up work or remedies will be given for losses of access to LEARN and academic consequences arising from administrative issues with the Registrar's Office.

3. Classroom Protocols
   - No private conversations in class will be tolerated.
   - Turn off all cell phone communication devices and put them away for the duration of the lecture.
   - Photographic devices are not permitted in class; this is a University regulation.
   - In order to maintain a learning environment in this classroom, I expect the highest level of courtesy and professional behaviour from every student.

University Statements

Economics Department Deferred Final Exam Policy

Deferred Final Exam Policy found at https://uwaterloo.ca/economics/undergraduate/resources-and-policies/deferred-final-exam-policy

Cross-listed course

Please note that a cross-listed course will count in all respective averages no matter under which rubric it has been taken. For example, a PHIL/PSCI cross-list will count in a Philosophy major average, even if the course was taken under the Political Science rubric.

Academic Integrity

Academic Integrity: In order to maintain a culture of academic integrity, members of the University of Waterloo are expected to promote honesty, trust, fairness, respect and responsibility. See the UWaterloo Academic Integritity webpage and the Arts Academic Integrity webpage for more information.

Discipline: A student is expected to know what constitutes academic integrity, to avoid committing academic offences, and to take responsibility for his/her actions. A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (e.g., plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course professor, academic advisor, or the Undergraduate Associate Dean. When misconduct has been found to have occurred, disciplinary penalties will be imposed under Policy 71 – Student Discipline. For information on categories of offenses and types of penalties, students should refer to Policy 71 - Student Discipline. For typical penalties check Guidelines for the Assessment of Penalties.

Grievance: A student who believes that a decision affecting some aspect of his/her university life has been unfair or unreasonable may have grounds for initiating a grievance. Read Policy 70 - Student Petitions and Grievances, Section 4. When in doubt, please be certain to contact the department's administrative assistant who will provide further assistance.

Appeals: A decision made or penalty imposed under Policy 70 - Student Petitions and Grievances (other than a petition) or Policy 71 - Student Discipline may be appealed if there is a ground. A student who believes he/she has a ground for an appeal should refer to Policy 72 - Student Appeals.

Accommodation for Students with Disabilities
Note for students with disabilities: The AccessAbility Services office, located on the first floor of the Needles Hall extension (1401), collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic

accommodations to lessen the impact of your disability, please register with the AS office at the beginning of each academic term.

Mental Health Support
All of us need a support system. The faculty and staff in Arts encourage students to seek out mental health supports if they are needed.

On Campus
- Counselling Services: counselling.services@uwaterloo.ca / 519-888-4567 xt 32655
- [MATES](): one-to-one peer support program offered by Federation of Students (FEDS) and Counselling Services
- Health Services Emergency service: located across the creek form Student Life Centre

Off campus, 24/7
- [Good2Talk](): Free confidential help line for post-secondary students. Phone: 1-866-925-5454
- Grand River Hospital: Emergency care for mental health crisis. Phone: 519-749-433 ext. 6880
- [Here 24/7](): Mental Health and Crisis Service Team. Phone: 1-844-437-3247
- [OK2BME](): set of support services for lesbian, gay, bisexual, transgender or questioning teens in Waterloo. Phone: 519-884-0000 extension 213

Full details can be found online at the Faculty of ARTS [website]()
Download [UWaterloo and regional mental health resources (PDF)]()
Download the [WatSafe]() app to your phone to quickly access mental health support information

Territorial Acknowledgement
We acknowledge that we are living and working on the traditional territory of the Attawandaron (also known as Neutral), Anishinaabe and Haudenosaunee peoples. The University of Waterloo is situated on the Haldimand Tract, the land promised to the Six Nations that includes six miles on each side of the Grand River.