Department of Economics
University of Southern California

Spring 2025
ECON 570: Big Data Econometrics

Course Information
Time and Location: Monday, Wednesday: 2:00 - 3:20 pm
Location: THH 202

Instructor: Dong Woo Hahm (dongwooh@usc.edu)
Office Hours: TBD

Teaching Assistant
TA: TBD
Office hours:
Sessions:
Email:

Course Overview
This course provides an introduction to machine learning and related methods for big data from the perspective of economics. Students will be introduced to modern estimation methods for high-dimensional data, which will be illustrated through applications to causal inference and prediction problems in economics, business, and related fields. Students will gain experience working with these methods through programming assignments. The course will be focused on methodology and its practical application and will culminate in an empirical project in which students apply course concepts to real-world data. By the end of the course, students should be able to do the following.
- Apply machine learning methods to estimate causal effects in experimental and observational settings and solve prediction problems.
- Understand the uses and limitations of machine learning for answering economic questions.
- Implement machine learning algorithms using R.
Econometrics at the level of ECON 513 is required, along with calculus at the level of MATH 226 and linear algebra at the level of MATH 225. Some programming background may be needed (see below).

Course Requirements

Grades will be based on problem assignments (40%), midterm quiz (10%), final quiz (10%), and a final project (40%).

- Assignment (40%): Regular programming exercises will be assigned to reinforce the concepts taught in class, as well as offer an opportunity for students to code and implement the algorithms covered. Students will work in groups of 4-5, which will be randomly assigned at the beginning of the course. While each student is expected to contribute, only one problem set will be turned in per group. R **must be** used as the language for programming exercises. Problem sets will be turned in through Brightspace, along with code and any graphs generated. Assessment will be based on whether the right approaches were used and whether the right solutions were obtained. Problem sets will be discussed in class a week or more after the original due date. The absolute deadline for turning in a problem set is when solutions are made available: homework turned in after this second deadline will receive zero points. 30% will be taken off for problem sets that are turned in late, but before the problem set is discussed in class or solutions are made available, whichever comes first.

- Midterm and final quizzes (10% each, 20% total): A short quiz will be held near the middle of the semester and toward the end of the semester. Basic knowledge of concepts covered during lectures will be tested.

- Final project (40%): Students are expected to work in groups of 4-5 students each[1] and apply course concepts to a real-world data set to tackle an empirical problem that interests them. Students are expected to collect their own data or use publicly available data as part of the project. Suggestions for starting points for publicly available data will be posted on Brightspace. Each group will submit a writeup, as well as code to reproduce the analysis. Groups will give short presentations of their work in class during the last few lectures. Assessment will be based on how appropriately the quantitative tools were applied. Due date for this project is TBA.

Optional Textbooks

Much of the lecture material draws from the first textbook below.

- [BDS] Taddy, M. (2019) *Business Data Science: Combining Machine Learning and Economics to Optimize, Automate, and Accelerate Business Decisions*. McGraw Hill.

- [ISL] Gareth, J.; Witten, D.; Hastie, T. and Tibshirani, R. (2017) *An Introduction to Statistical Learning*. Springer.

- [MM] Angrist, J. and Pischke, J. (2014) *Mastering Metrics: The Path from Cause to Effect*. Princeton University Press.

Software

---

[1] You may form your group on your own. I will ask you to submit the list of group members around the middle of the course.

USC

This course will use R, a free statistical programming language. While you do not need to have prior knowledge of R to take this course, you will be expected to learn R on your own time not only through lectures and homework, but also by searching independently for relevant commands, tutorials and documentation using Google, StackOverflow, the R help() command, etc. Programming experience in another language should suffice, so long as you are willing to learn R on your own time.

Course Outline *(Note: subject to change throughout the semester.)*
1. Introduction (0.5 weeks; BDS Intro; ISL Ch.2)
    a. Prediction v.s. causal inference
    b. Intro to R
2. Sampling (1 week; BDS Ch. 1; ISL Ch. 5.2).
    a. CLT and standard errors.
    b. Bootstrap.
3. Regression (1.5 weeks; BDS Ch. 2; ISL Ch. 3).
    a. Linear conditional mean model.
    b. Logistic regression.
4. Fundamental concepts in ML (2.5 weeks; BDS Ch. 3; ISL Ch. 2, 5.1, 6).
    a. Cross-validation.
    b. Regularization.
    c. Bias-variance trade-off.
5. Classification (1.5 weeks; BDS Ch. 4; ISL Ch. 4.3).
    a. Logistic lasso.
    b. Multinomial logit.
6. Causal inference (2.5 weeks; BDS Ch. 5, 6; MM Ch. 1, 2).
    a. Potential outcomes framework
    b. Randomized control trials.
    c. High-dimensional controls.
7. Woodlands (1 week; BDS Ch. 9; ISL Ch. 8).
    a. Classification and regression trees.
    b. Random forests.
8. Unsupervised learning (2 weeks; BDS Ch. 7; ISL Ch. 10).
    a. K-means clustering
    b. Principal components analysis.
    c. Partial least squares.
9. Natural language processing (1.5 weeks; BDS Ch. 8).
    a. Tokenization and bag-of-words.
    b. Topic models.
10. Artificial intelligence (1 week; BDS Ch. 10).
    a. Overview of AI systems
    b. Deep neural networks.
    c. Stochastic gradient descent.

Statement on Academic Conduct and Support Systems

Academic Integrity:

The University of Southern California is a learning community committed to developing successful scholars and researchers dedicated to the pursuit of knowledge and the dissemination of ideas. Academic misconduct, which includes any act of dishonesty in the production or submission of academic work, compromises the integrity of the person who commits the act and can impugn the perceived integrity of the entire university community. It stands in opposition to the university's mission to research, educate, and contribute productively to our community and the world.

All students are expected to submit assignments that represent their own original work, and that have been prepared specifically for the course or section for which they have been submitted. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s).

Other violations of academic integrity include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), collusion, knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

The impact of academic dishonesty is far-reaching and is considered a serious offense against the university. All incidences of academic misconduct will be reported to the Office of Academic Integrity and could result in outcomes such as failure on the assignment, failure in the course, suspension, or even expulsion from the university.

For more information about academic integrity see the student handbook or the Office of Academic Integrity's website, and university policies on Research and Scholarship Misconduct.

Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Support Systems:

[Counseling and Mental Health](#) - *(213) 740-9355 – 24/7 on call*

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

[988 Suicide and Crisis Lifeline](#) - *988 for both calls and text messages – 24/7 on call*

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline is comprised of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

[Relationship and Sexual Violence Prevention Services (RSVP)](#) - *(213) 740-9355(WELL) – 24/7 on call*

Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

[Office for Equity, Equal Opportunity, and Title IX (EEO-TIX)](#) - *(213) 740-5086*

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

[Reporting Incidents of Bias or Harassment](#) - *(213) 740-5086 or (213) 821-8298*

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

[The Office of Student Accessibility Services (OSAS)](#) - *(213) 740-0776*

OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

[USC Campus Support and Intervention](#) - *(213) 740-0411*

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

[Diversity, Equity and Inclusion](#) - *(213) 740-2101*

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - *UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call*

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - *UPC: (213) 740-6000, HSC: (323) 442-1200 – 24/7 on call*

Non-emergency assistance or information.

Office of the Ombuds - *(213) 821-9556 (UPC) / (323-442-0382 (HSC)*

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Occupational Therapy Faculty Practice - *(323) 442-2850 or* otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.