ECON 178: Economic and Business Forecasting, Winter 2020

Ying Zhu

- Lecture. Section A: TuTh 11:00am 12:20pm WLH. Section B: TuTh 12:30pm 1:50pm WLH
- Discussion. Section A: M 3:00pm 3:50pm WLH. Section B: M 4:00pm 4:50pm WLH
- Instructor. Ying Zhu; Email: yiz012@ucsd.edu; Office Hours: Tuesday 2:30pm 3:30pm, Economics Department, #321
- **Teaching assistants.** Davide Viviano (dviviano@ucsd.edu). Julian Martinez Iriarte (jum024@ucsd.edu). TA Office Hours: To be decided
- Reader. Connor Goldstick (cgoldsti@ucsd.edu)

Course Description

This course provides an introduction to a number of modern statistical learning methods, including stepwise selection, ridge regression, the Lasso, trees, SVM, and Gaussian processes. We will discuss the basic concepts and computational issues behind these methods, the concepts of prediction, uncertainty quantification and causal inference, as well as applications in economics and business.

Textbook and Other Course Materials

- Required: An Introduction to Statistical Learning with Applications in R, 7th edition (James, Witten, Hastie, Tibshirani): http://faculty.marshall.usc.edu/gareth-james/ISL/
- Online resources: https://www.r-bloggers.com/in-depth-introduction-to-machine-learning-in-15-hours-of-expert-videos/
- My lecture notes (motivations, concepts, methods, theory, computational issues)
- Discussion notes (reviewing lectures, demonstrations in R)¹

Prerequisites

ECON 120C

¹Thank to Davide Viviano who has prepared these notes.

Software

This course is heavily based on the statistical software R. Each problem set will include empirical questions based on real data sets. To answer the empirical questions, you will need to use R. Your teaching assistants will run tutorials on R and explain the necessary commands needed to complete the problem sets. This will be a good opportunity for you to be familiar with this software and see how it works in practice.

Grades

- 1. Bi-weekly problem sets 45%. Late problem sets will not be accepted except with my prior consent or in unusual circumstances permitted by University policies. The problem sets will include both conceptual and applied questions. The purpose of the problem sets is to help you learn the material and assess your progress.
- 2. A final project 55%. You will be provided with a data set and guidelines that help you to work through the project.
- 3. Bonus opportunity 20%. Competition for top predictors (to be discussed in the lectures).

Academic Dishonesty

Academic dishonesty will be dealt with according to the University policies (https://academicintegrity.ucsd.edu/). Cheating includes, but is **not limited to**: copying someone else's homework, stealing someone else's homework, etc. **NOTE: You are allowed to work together on *Bi-weekly problem Sets*. However, each of you must hand in your OWN versions of the answers. You are not allowed to work together on the final project and the bonus opportunity.** Penalties on cheating include, among others, a failing grade for the course.

Students with Disabilities

The Americans with Disabilities Act of 1990 prohibits discrimination based on disability. Students with disabilities are encouraged to discuss their concerns with the instructor, preferably at the beginning of the quarter. University-wide resources can be found at https://students.ucsd.edu/well-being/disability-services/index.html.

Course Outline (Tentative)

- 1. Motivation and basic concepts of statistical learning
- 2. Review of the linear regression model and least squares estimator
- 3. Ridge regressions and the Bayesian interpretation
- 4. Feature selections: best subset selection, stepwise selection, the Lasso

5. Tree methods

- 6. SVM
- 7. Uncertainty quantification and Gaussian process
- 8. Causal inference