

EC524: Prediction and machine learning

Following your first course on econometrics and causal inference, this course turns to examining the tools available for prediction. Put simply, we are now focusing on \hat{y} rather than $\hat{\beta}$ from the model $y = \alpha + \beta x$.

Books

Each book (except one of the recommended books) is available for free online.

The physical copies are also very reasonably priced.

Required

- [The Hundred-Page Machine Learning Book](#) (100ML)
- [Introduction to Statistical Learning](#) (ISL)
- [Data Visualization](#) (Data Viz)

Recommended

- [R for Data Science](#)
- [Introduction to Data Science](#) (not available without purchase)
- [The Elements of Statistical Learning](#) (ESL, the big brother of ISL)

Topics

0. Prediction and *machine learning*

- 0. Prediction vs. causal inference **Readings** [Prediction Policy Problems](#) by Kleinberg *et al.* (2015)

1. Exploratory data analysis

- 0. Building insights from graphics **Readings** [Data Viz](#) Preface, Ch1
- 1. Learning ggplot2 **Readings** [Data Viz](#) Ch3

2. Supervised learning

- 0. An introduction to machine learning **Readings** 100ML Preface, Ch1–Ch4; ISL 2.1–2.2
- 1. LASSO and Ridge regression **Readings** ISL 6.1–6.3, 6.6
- 2. Classification trees **Readings** 100ML 3.3; ISL 8.1
- 3. **Aside** Resampling methods and other best practices **Readings** 100ML Ch5; ISL Ch5
- 4. Regression trees **Readings** 100ML 3.3; ISL 8.1
- 5. SVM **Readings** 100ML 3.4; ISL 9.1–9.4
- 6. Neural nets **Readings** 100ML 6
- 7. Boosting and ensembles **Readings** 100ML 7.5 and Ch8
- 8. Random forests **Readings** ISL
- 9. Additional topics **Readings** 100ML Ch7 and Ch11

3. Unsupervised learning

- 0. Introduction to unsupervised learning **Readings** 100ML Ch9; ISL 10.1
- 1. Principal components analysis **Readings** ISL 10.2; 100ML 9.3
- 2. Nearest-neighbor matching, K-means, and hierarchical clustering **Readings** 100ML Ch9; ISL 10.3