Machine Learning for Empirical Economic Research, Part I

Andreas Dzemski

June 5, 2019

Course literature

The textbook for this course is

Trevor Hastie, Robert Tibshirani, and Jerome Friedman (2009). *The elements of statistical learning: data mining, inference, and prediction, Springer Series in Statistics.* Springer New York.

It is available for free online at `https://web.stanford.edu/~hastie/ElemStatLearn/`.

Examination

PhD students can study this course (i.e., both Part 1 and Part 2) as a reading course (individuellt tagande) and receive 3 ECTS points. To receive credit, students have to study the reading list, participate in the discussions in class and complete a written assignment. The written assignment will be customized to the students individual research interests. It can be either a research proposal for an empirical project using a Machine Learning method, or an extended referee report for an applied or theoretical paper on Machine Learning methods in economics.

If you need inspiration for your project, conferences and workshop programs are good way to see where the literature is heading. For example, `https://www.barcelonagse.eu/summer-forum/workshop-machine-learning-economics`.

Lectures

Lecture 1

June 4, 3-5pm in F45

We will try to define what Machine Learning is and discuss how it is different from more traditional econometric methods and how it can be part of an empirical research strategy. Moreover, we will discuss tools (programming languages, libraries) useful for implementing Machine Learning methods.

Reading: Mullainathan and Spiess (2017); Varian (2014) Chapter 7.1-7.3, 7.10, 7.12 of Hastie, Tibshirani, and Friedman (2009)

Lecture 2

June 5, 3-5pm in C32

We introduce the supervised learning framework and some central concepts related to it such as regularization and k-fold cross-validation. We discuss penalized linear regression (Ridge and Lasso) as an example of a supervised learning method.

Reading: Chapter 2.1-2.7, 2.9, 3.4 of Hastie, Tibshirani, and Friedman (2009)

Lecture 3

June 7, 3-5pm in B21

We apply high-dimensional regression techniques to estimate treatment effects.

Reading: Belloni, Chernozhukov, and Hansen (2014), Belloni, Chernozhukov, Fernández-Val, et al. (2017), and Athey, Imbens, and Wager (2018)

References

Athey, Susan (2017). "Beyond prediction: Using big data for policy problems". In: *Science* 355.6324, pp. 483–485.

Athey, Susan, Guido W Imbens, and Stefan Wager (2018). "Approximate residual balancing: debiased inference of average treatment effects in high dimensions". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 80.4, pp. 597–623.

Becerra, Victor M, Roberto KH Galvão, and Magda Abou-Seada (2005). "Neural and wavelet network models for financial distress classification". In: *Data Mining and Knowledge Discovery* 11.1, pp. 35–55.

Belloni, Alexandre, Victor Chernozhukov, Ivan Fernández-Val, et al. (2017). "Program evaluation and causal inference with high-dimensional data". In: *Econometrica* 85.1, pp. 233–298.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen (2014). "Inference on treatment effects after selection among high-dimensional controls". In: *The Review of Economic Studies* 81.2, pp. 608–650.

Blumenstock, Joshua, Gabriel Cadamuro, and Robert On (2015). "Predicting poverty and wealth from mobile phone metadata". In: *Science* 350.6264, pp. 1073–1076.

Cassel, Claes M, Carl E Särndal, and Jan H Wretman (1976). "Some results on generalized difference estimation and generalized regression estimation for finite populations". In: *Biometrika* 63.3, pp. 615–620.

Gentzkow, Matthew and Jesse M Shapiro (2010). "What drives media slant? Evidence from US daily newspapers". In: *Econometrica* 78.1, pp. 35–71.

Graham, Bryan S, Cristine Campos de Xavier Pinto, and Daniel Egel (2012). "Inverse probability tilting for moment condition models with missing data". In: *The Review of Economic Studies* 79.3, pp. 1053–1079.

Hartford, Jason et al. (2017). "Deep IV: A flexible approach for counterfactual prediction". In: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, pp. 1414–1423.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009). *The elements of statistical learning: data mining, inference, and prediction, Springer Series in Statistics*. Springer New York.

Jean, Neal et al. (2016). "Combining satellite imagery and machine learning to predict poverty". In: *Science* 353.6301, pp. 790–794.

Kogan, Shimon et al. (2009). "Predicting risk from financial reports with regression". In: *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, pp. 272–280.

Lee, Jason D et al. (2016). "Exact post-selection inference, with application to the lasso". In: *The Annals of Statistics* 44.3, pp. 907–927.

Mullainathan, Sendhil and Jann Spiess (2017). "Machine learning: an applied econometric approach". In: *Journal of Economic Perspectives* 31.2, pp. 87–106.

Newey, Whitney K (1997). "Convergence rates and asymptotic normality for series estimators". In: *Journal of econometrics* 79.1, pp. 147–168.

Newey, Whitney K and James L Powell (2003). "Instrumental variable estimation of nonparametric models". In: *Econometrica* 71.5, pp. 1565–1578.

Robins, James (1986). "A new approach to causal inference in mortality studies with a sustained exposure period". In: *Mathematical modelling* 7.9-12, pp. 1393–1512.

Robins, James M and Ya'acov Ritov (1997). "Toward a curse of dimensionality appropriate (CODA) asymptotic theory for semi-parametric models". In: *Statistics in medicine* 16.3, pp. 285–319.

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1, pp. 267–288.

Varian, Hal R (2014). "Big data: New tricks for econometrics". In: *Journal of Economic Perspectives* 28.2, pp. 3–28.