

Das Institut für Volkswirtschaftslehre und Ökonometrie weist auf die folgende, aus Studienzuschüssen finanzierte Blockvorlesung auf Masterniveau hin:

Machine Learning

Prof. Dr. Martin Spindler, Universität Hamburg

Inhalt der Veranstaltung. The course will provide a practical introduction to modern high-dimensional function fitting methods — also known as machine learning (ML) methods — for efficient estimation and inference on treatment effects and structural parameters in empirical economic models. Participants will use R to allow them to immediately internalize and use the techniques in their own academic and industry work. All lectures, except the introductory one, will be accompanied by R-code that can be used to reproduce the empirical examples in the lectures. Thus, there will be no gap between theory and practice.

Gliederung. Siehe Anlage.

Voraussetzungen. Fundierte Kenntnisse in Ökonometrie und Statistik sowie der Software R, z. B. durch erfolgreichen Besuch der Master-Veranstaltung Methoden der Ökonometrie.

Unterrichtssprache. Englisch.

Ort und Zeit. Die Lehrveranstaltung findet **geblockt vom 11.03.2019 bis 13.03.2019 (jeweils 9:00 – ca. 18:00 Uhr)** im Hörsaal H 26 (Vielberth-Gebäude) sowie teilweise auch im CIP-Pool RWL U01 (rechts) statt.

Klausurtermine. Die 60-minütige Klausur findet statt am 20.03.2019, 10.30 bis 11.30 Uhr im Hörsaal H 8. Die Wiederholungsklausur findet statt am 24.04.2019, 14.30 bis 15.30 Uhr. Der Hörsaal wird noch bekanntgegeben. Die Wiederholungsklausur ist **nicht** für Erstschreiber geöffnet.

Klausuranmeldung. Die Anmeldung zu den Klausuren erfolgt über FlexNow. Der An- bzw. Abmeldezeitraum für die Klausur ist vom 27.02.2019 bis 13.03.2019. Die Wiederholungsklausur ist wie gesagt nicht für Erstschreiber offen. Der An- bzw. Abmeldezeitraum für die Wiederholungsklausur ist 03.04.2019 bis 17.04.2019.

Anrechnung. Diese Master-Lehrveranstaltung des Instituts für Volkswirtschaftslehre und Ökonometrie kann mit 6 Kreditpunkten (ECTS) in das Wahlmodul oder in die Schwerpunktmodulgruppen „Empirische Wirtschaftsforschung“ oder „Finanzmärkte“

eingebraucht werden (wobei Sie diese Entscheidung bei der Anmeldung in Flexnow treffen). Bachelorstudierende können diese Master-Lehrveranstaltung in der zweiten Studienphase in das Wahlmodul oder die obengenannten Schwerpunktmodulgruppen einbringen. Die Veranstaltung wird für das Wintersemester 2018/19 gewertet.

WICHTIG: Anmeldung und Vorlesungsunterlagen. Bitte registrieren Sie sich bereits im Vorfeld für die Veranstaltung in GRIPS unter <https://elearning.uni-regensburg.de/course/view.php?id=23194>. Dort werden im Vorfeld auch die Vorlesungsunterlagen zur Verfügung gestellt.

Für Rückfragen wenden Sie sich bitte an Prof. Dr. Rolf Tschernig (rolf.tschernig@ur.de).

Martin Spindler (Instructor)

Machine Learning

The Course Idea. The course will provide a practical introduction to modern *high-dimensional* function fitting methods — also known as machine learning (ML) methods — for efficient estimation and inference on treatment effects and structural parameters in empirical economic models. Participants will use R to allow them to immediately internalize and use the techniques in their own academic and industry work. All lectures, except the introductory one, will be accompanied by R-code that can be used to reproduce the empirical examples in the lectures. Thus, there will be no gap between theory and practice.

Outline

0. Recap: Linear Regression

1. Understanding the Inference Strategy via Partialling Out and Adaptivity

- a. Theory: Frisch-Waugh Partialling Out. Adaptivity.
- b. Laying out a strategy for the use of non-sparse and generic ML methods.
- c. **R Practicum: Mincer Equations, Barro-Lee, and Acemoglu-Johnson-Robinson examples.**
- d. References: 3, 4, 6.

2. ML Methods for Prediction

- a. Penalization Regression Methods: Ridge, Lasso, Elastic Nets, etc.
- b. Regression Trees, Random Forests, Boosted Trees.
- c. Modern Nonlinear Regression via Neural Nets and Deep Learning.
- d. Aggregation of the ML methods.
- e. **R Practicum: Simulated, Wage, and Pricing Examples.**
- f. References: 1 and 2.

3. ML Methods for Causal Parameters — “Double” Machine Learning for Causal Parameters in Treatment Effect Models and other Econometric Models

- a. Using generic ML (beyond Lasso) to Estimate Coefficients in Partially Linear Methods
- b. Using generic ML to estimate ATE, ATT, LATE in Heterogeneous Treatment Effect Models
- c. Using generic ML methods to estimate structural parameters in GMM problems.
- d. **R Practicum: 401(k) Example.**
- e. References: 5, 6, 7, 8.

4. Scalability: Working with Large Data. MapReduce, Hadoop and all that

Requirements

Please bring your computer to class. Install R and R-studio. Install packages “hdm”, “glmnet”, “nnet”, “randomForest”, “rpart”, “rpart.plot”, “gbm” from cran (e.g. type `install.packages(“gbm”)`). If you are not familiar with R, try out several introductory tutorials that are available online.

References

1. [The Elements of Statistical Learning](#) 2009, 2nd ed. by T. Hastie, R. Tibshirani, and J. Freedman. The book can be downloaded for free!
2. [An Introduction to Statistical Learning with Applications in R](#), by G. James, D. Witten, T. Hastie and R. Tibshirani. The website has a lot of handy resources.
3. ["High-Dimensional Methods and Inference on Treatment and Structural Effects in Economics"](#), *Journal of Economic Perspectives* 2014, A. Belloni, V. Chernozhukov, C. Hansen, R code implementation is in package “hdm”
4. ["Inference on Treatment Effects After Selection Amongst High-Dimensional Controls \(with an Application to Abortion and Crime\)"](#), *Review of Economic Studies* 2013, A. Belloni, V. Chernozhukov, C. Hansen, Stata and Matlab programs are [here](#); replication files [here](#). R code implementation in package “hdm”.
5. ["Robust Inference on Average Treatment Effects with Possibly More Covariates than Observations"](#), *Journal of Econometrics* 2015, M. Farrell.
6. ["Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach"](#), *Annual Review of Economics* 2015, V. Chernozhukov, C. Hansen, and M. Spindler. R code implementation in package “hdm”.
7. ["Program Evaluation and Causal Inference with High-Dimensional Data"](#), *Econometrica* 2016, A. Belloni, V. Chernozhukov, I. Fernández-Val, C. Hansen, R code implementation in package “hdm”. Replication files via Econometrica website.
8. ["Double Machine Learning for Causal and Treatment Effects"](#), V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey.
9. ["Big Data: New Tricks for Econometrics,"](#) *Journal of Economic Perspectives* 2014, H. Varian.
10. ["Economics in the age of big data,"](#) *Science* 2014, L. Einav, J. Levin.
11. ["Prediction Policy Problems,"](#) *American Economic Review P&P* 2015, J. Kleinberg, J. Ludwig, S. Mullainathan, Z. Obermeyer.