

Prof. Melanie Krause, Ph.D.
Junior Professor
for Macroeconomics
University of Hamburg
Von-Melle-Park 5, Room 3103
20146 Hamburg
Tel: +49 40 42838 2082
Email: melanie.krause@uni-hamburg.de



Machine Learning Methods in Economics

Master's Elective in the Winter Semester 2020/21,
Course-Number 22-30.103

Course Description

Machine learning (ML) has become a buzzword of our time. With the amounts of data available growing exponentially, there is an increased demand to make use of all this information. In this course we take a look at the main methods of supervised and unsupervised learning from an economist's point of view. In particular, we will (i) learn about the most widely-used ML techniques such as variance-reduction techniques, neural networks and random forests, (ii) apply them to real-world data, (iii) discuss how ML-based prediction works differently from econometric inference.

Course Goals

Course participants will get to know the most important ML methods which are currently in use. The focus is on how ML methods can complement the statistical toolkit of an economist for research and applications. Course participants will learn how to apply a range of ML methods in Python, thus becoming well-equipped to conduct their own empirical project.

Course Settings

Lectures with Tutorials included Lectures with Tutorials included will be uploaded as videos to OpenOLAT. There will be weekly Zoom sessions to discuss the material (Mondays 11:00-12:00):

02.11., 09.11., 16.11., 23.11., 30.12., 07.12., 14.12., 04.01, 11.01., 18.01., 25.01., 01.02., 08.02., 15.02.

Course Language

The course language is English.

Course Prerequisites

Course participants are assumed to have a solid background in math, statistics and econometrics from their undergraduate studies. Furthermore, knowledge of the course material of the first semester master's course 'Applied Econometrics' is required.

Grading

In order to pass the course, students must (a) pass a 60-minute exam and (b) complete a research project involving ML methods learned in the course. They should suggest their own research question and data set, explaining why ML is particularly suited in this situation. They should then carry out the research project and discuss their results.

The exam will make up 50% of the final grade and the project 50%.

Exam Dates:

First Date: Monday, 22 February 2021, 15:15-16:15

Second Date: Thursday, 25 March 2021, 09:15-10:15

Course participants have to send their completed projects to melanie.krause@uni-hamburg.de by 15 March 2021 (23:59).

Course Literature

Main Course Books and Papers:

- Hastie, Tibsharini and Friedman (2017), "The Elements of Statistical Learning - Data Mining, Inference, and Prediction", Second Edition. ('The Bible of ML' - sometimes going far beyond this course)
- Mueller and Guido (2017), "Introduction to Machine Learning with Python", First Edition. (For the applications in Python)
- Athey and Imbens (2019), "Machine Learning Methods Economists Should Know About", Working Paper at Arxiv (<https://arxiv.org/pdf/1903.10075.pdf>)

Additional Books:

- Kuhn and Johnson (2012), "Applied Predictive Modeling", First Edition. (Overview of key concepts, but less technical, many applications in R)
- Varoquaux, Gouillart, Vahtras, de Bruyl (2020), "Scipy Lecture Notes", Available on www.scipy-lectures.org (Extensive Python Tutorial for scientific analysis)
- Heinold (2012), "A Practical Introduction to Python Programming", First Edition. (More Python with many practical examples, not ML-specific)

Course Structure (preliminary)

Topics discussed will include:

- Econometrics vs Machine Learning in Economics
- Supervised vs Unsupervised Learning
- Introduction to Python
- Linear Regression
- Logistic Regression
- The Bias-Variance Tradeoff
- Ridge and LASSO Regression
- Decision Trees
- Random Forest
- Gradient Boosted Trees
- Support Vector Machines
- Neural Networks
- Pre-Processing Data and Feature Selection
- Cross-Validation
- Model Performance and Model Selection
- Principal Component Analysis
- Clustering
- Working with Text Data

Version as of 8 September 2020