Dr. Arne Jonas Warnke
DataCamp

Heidelberg University
Winter Term 2022/23

Machine Learning in Practice

This interactive course will discuss the application of machine learning and econometric models in practice. We will discuss both theory and best practices of modeling in business, production, and research. Thereby, we will not restrict ourselves to the core machine learning development. We will discuss how models are deployed and operated, the collaboration of modern data teams, and how to generate value based on machine learning models. The course will include self-learning elements between the lectures, and students will work on a project during the semester. The course will also give career advice to students interested in becoming data scientists.

Dates and room:

- October 20, 2022, 9 - 11 am, room 02.036 (AWI, Bergheimer Str. 58)
- November 3, 2022, 9 - 11 am, room 02.036 (AWI, Bergheimer Str. 58)
- November 17, 2022, 9 - 11 am, room 02.036 (AWI, Bergheimer Str. 58)
- December 15, 2022, 9 - 11 am, room 02.036 (AWI, Bergheimer Str. 58)
- January 12, 2022, 9 - 11 am, room 02.036 (AWI, Bergheimer Str. 58)

Example topics:

- The importance of (economic) theory in data science.
- Successful feature transformation (normalization, standardization, binning/bucketing)
- The ongoing success of linear models.
- What is the XGBoost algorithm, and why is it so popular?
- What does a modern machine learning team consist of and how do they collaborate (and what is the job role of a data scientist)?
- Train, validation, and test? When and why do I need two holdout datasets?
- How to ensure high accuracy of deployed models?

Grading consists of two components:

- Take-home project
- Final oral exam

Dr. Arne Jonas Warnke
DataCamp

Heidelberg University
Winter Term 2022/23

Prerequisites:

- Course Advanced Econometrics
- Statistics Knowledge
- Interest in Data Science and Machine Learning (Theory and Practice)

Contact and registration:

Please register via arne.warnke@gmail.com.

Literature:

Athey, Susan, and Guido W. Imbens. "Machine learning methods that economists should know about." *Annual Review of Economics* 11 (2019): 685-725.

Hastie, Trevor, Robert Tibshirani, Jerome H. Friedman, and Jerome H. Friedman. "*The elements of statistical learning: data mining, inference, and prediction*." Vol. 2. New York: springer, 2009.

Kreuzberger, Dominik, Niklas Kühl, and Sebastian Hirschl. "Machine Learning Operations (MLOps): Overview, Definition, and Architecture." *arXiv preprint arXiv:2205.02302* (2022).

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. "*An introduction to statistical learning*." Vol. 112. New York: springer, 2013.

Lakshmanan, Valliappa, Sara Robinson, and Michael Munn. *Machine learning design patterns*. O'Reilly Media, 2020.

Sculley, David, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-Francois Crespo, and Dan Dennison. "Hidden technical debt in machine learning systems." *Advances in neural information processing systems* 28 (2015).