Discipline: [Methods/Field]

Methods

1. Language

English

2. Title

Causal Machine Learning

3. Lecturer

Prof. Dr. Martin Spindler, Fakultät für Betriebswirtschaft, Universität Hamburg
Martin.spindler@uni-hamburg.de

4. Date and Location

06. – 09. Oktober 2025

Universität Hamburg
Moorweidenstr. 18
20148 Hamburg

The course will be offered over a four-day period comprising lecture, tutorial, and discussion sessions.

5. Course Description

5.1 Abstract and Learning Objectives

While AI and Machine Learning are mainly tailored for predictions, based on correlations, many important questions in industry and research are causal questions. Examples are pricing, marketing mix modelling, resource allocation, to name a few, or many questions in Corporate Finance, Human Resources or Management in general. The emerging field of Causal AI / ML combines causal inference with modern methods in machine learning as complex, to estimate causal effects in high-dimensional, complex data. Due to the rise of digitization, such data sets are more and more available and can be utilized for research.

Goal: The participants will learn the fundamental concepts and methods of Causal Machine Learning, in particular the Double Machine Learning approach, and will be able to apply the methods in their empirical research.

5.2 Content

1. Causal Inference via Randomized Experiments

2. Statistical Inference on Predictive Effects in High-Dimensional Linear Regression Models

3. Causal Inference via Conditional Ignorability

4. Causal Inference via Linear Structural Equations

5. Causal Inference via Directed Acyclical Graphs and Nonlinear Structural Equation Models

6. Valid Adjustment Sets from DAGs

7. Statistical Inference on Predictive and Causal Effects in Modern Nonlinear Regression Models

8. Advanced Unobserved Confounders, Instrumental Variables, and Proxy Controls

9. DML for IV and Proxy Controls Models and Robust DML Inference under Weak Identification

10. Statistical Inference on Heterogeneous Treatment Effects

11. Difference-in-Differences

12. Regression Discontinuity Designs

13. Outlook and Recent Developments

5.3 Schedule (including start and end time)

The course consists of several lecture (L) and hands-on session (H) sessions.

Pre-course stage:

1) Study papers from reading list
2) Familiarize with Python and Jupyter notebooks

Post-course stage (4 to 6 weeks):
Development of a Jupyter notebook demonstrating the use of Causal ML in research. Specific tasks will be agreed with participants and should ideally display a strong link to the participant's Ph.D. topic.

Course stage

Day	Time	Lecture	Hands-on Session
Day 1	9-17:30	L1-L4	1 hands-on seession (90 min)
Day 2	9-17:30	L5-L8	1 hands-on seession (90 min)
Day 3	9-17:30	L8-10	1 hands-on seession (90 min)
Day 4	9-15:30	L11-L13	1 hands-on seession (90 min)

5.4 Course format

Vorlesung und Tutorials mit Hands-on Examples

6. Preparation and Literature

6.1 Prerequisites

Linear Regression and Statistics
Machine Learning

6.2 Essential Reading Material

Chernozhukov, Victor; Hansen, Chris; Kallus, Nathan; Spindler, Martin Syrgkanis, Vasilis (2023): Applied Causal Inference Powered by ML and AI. Book draft in preparation, causalml-book.org.

Huber, Martin (2023): Causal Analysis. MIT Press.

6.3 Additional Reading Material

James, Witten, Hastie, Tibshirani: Introduction to Statistical Learning, Springer, https://www.statlearning.com/

6.4 To prepare

Participants are expected to study the essential reading material. Familiarity with literature from the additional reading material list is beneficial. The Ph.D. course “Machine Learning”, which is also offered in the VHB ProDok lecture series, provides an excellent foundation for the course. To prepare for the practical exercises and course assignment, participants are required to familiarize themselves with the Python programming language and Jupyter notebooks. To that end, participants might find the following textbook useful:

- Géron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. 2nd Edition. O’Reilly Media Inc.
- VanderPlas, J. (2016). Python Data Science Handbook: Essential Tools for Working with Data. Sebastopol, CA, USA: O'Reilly Media. https://jakevdp.github.io/PythonDataScienceHandbook/

7. Administration

7.1 Max. number of participants

20 participants

7.2 Assignments

none

7.3 Exam

After the course, participants are required to complete a machine learning assignment and write-up results in the form of a computational essay (i.e., Jupyter Notebook). Typically, each participant will work on a different modeling task. Ideally, the assignment task connects to a research project that the participant is involved. Alternative assignment topics include the replication of a published causal machine learning paper or working on a Kaggle competition (http://www.kaggle.com). The schedule of

the course leaves room for discussing possible topics for the assignment. Student will submit their solution to the assignment roughly six weeks after the end of the course period. The submitted notebooks will be graded according to the quality of the exposition, the complexity of the modeling tasks, and the degree to which machine learning concepts have been used successfully.

7.4 Credits

The course corresponds to a scope of 6 LP/ECTS

8. Working Hours

Working Hours	**hours**
Mandatory readings	20h
Preparation for programming part / study of pre-course Jupyter notebooks	40h
Active participation in class	30h
Final exam (practical assignment to be completed and written-up after the course)	90h
SUM	**180 h**