Master's Seminar in Economics

Statistical Learning in Economics

Winter Semester 2026/2027

Technische Universität Braunschweig

Institut für Volkswirtschaftslehre
Institute of Economics

Content

Statistical Learning (SL) combines the fields of Econometrics and Data Science in a manner that yields new and deep insights into data, allows for more precise predictions and takes the statistical properties of explanatory variables into account. Thus, SL methods can be applied in numerous scientific topics among others economic forecasting, individual level marketing campaigns or estimation of heterogeneous treatment effects. Students that are willing to attend this course must have basic knowledge on statistics, programming in Python or R as well as economic theory.

The focus of the course will be the application to financial and economic data analysis. The theory part will take place in the first half of the course and will cover the most relevant Statistical Learning methods in economics. The application part of the course will then follow in the format of programming exercises, where students will train their practical skills and apply Statistical Learning methods to economic and financial datasets. At the end of the course, students will conduct their own Statistical Learning analysis and prepare a term paper based on this analysis.

Students are explicitly encouraged to use generative AI (such as Claude, ChatGPT, or similar tools) for all project steps including literature research, data exploration, code development, and paper writing. However, all AI interactions must be documented and submitted as part of the course deliverables. The number of tokens consumed across all AI tools will be incorporated as an efficiency metric into the grading rubric. This approach emphasizes developing competence in directing AI effectively rather than pure technical implementation.

Prerequisites

Basic knowledge of Microeconomics, Macroeconomics and Econometrics is required. Prior knowledge of Machine Learning or Data Analysis is helpful but not required.

Registration

Students can register for the seminar via Stud.IP from 01.09.2026 at 08:00. Registration via TUconnect is mandatory.

Important dates

- from 21.09.2026 to 08.10.2026: Registration via Stud.IP
- 09.10.2026: Kick-off and theory crash-course (9:00-12:00, VWL institute)
- from 22.10.2026 to 05.11.2026: Registration via TUConnect (binding!)
- 23.10.2026: Application R-Lab 01 (13:00-16:00, online)
- 30.10.2026: Application R-Lab 02 (13:00-16:00, online)
- 04.11.2026: Application R-Lab 03 (9:00-12:00, online)
- 04.12.2026: Consultation (10:00-12:00, online)
- 29.01.2027: Submission deadline for the term paper, AI chat logs, and token documentation (17:00, via E-mail)
- 05.02.2027: Block seminar, presentation (09:00-12:00, VWL institute)

Course structure

1. Theory

- Introduction and Statistical Learning Trade-offs: Curse of Dimensionality, Bias-Variance Trade-off, Overfitting vs. Underfitting
- Overcome the Trade-offs: Cross-Validation, Bootstrapping
- Linear Regression: Simple Linear, Multiple Linear, Extensions
- Model Selection and Regularization: Stepwise Selection, Shrinkage Methods
- Classification: Logistic Regression, Discriminant Analysis
- Trees and SVM: Bagging, Boosting, SVM Classifier

2. Application

- R-Lab 01 Introduction to R and Linear Regression: Loading and Inspecting Data in R; Linear Regression; Non-linear Transformation and Qualitative Information
- R-Lab 02 Classification and Resampling Methods: Logistic Regression; Linear Discriminant Analysis and KNN; Validation Set Approach and Cross-Validation; Bootstrapping
- R-Lab 03 Model Selection/Regularization and Decision Trees/SVM: Best Subset Selection + Selection Using Validation Set and CV; Ridge Regression and the Lasso; Classification and Regression Trees; Bagging and Boosting; Support Vector Classifier and Machine

Examination

The following are mandatory for all seminar participants:

- Attendance at the kick-off meeting/theory crash-course, and application labs
- Active participation and contribution to seminar presentation
- Submission of the term paper

Unlike previous semesters, students are responsible for identifying and formulating their own research question. The research question should be:

- Grounded in economic or financial theory
- Addressable using statistical learning methods
- Supported by publicly accessible or university-provided datasets

The following components constitute the term paper and grading basis:

- Literature review and theoretical foundation for the chosen research question
- Data description and summary statistics
- Empirical methodology (choice and justification of SL methods)
- Results and interpretation
- R code and reproducible analysis

All interactions with generative AI tools must be exported and submitted alongside the term paper. This should include:

- Complete conversation transcripts (in .txt, .pdf, or .json format)
- Clear labeling of the AI tool used (e.g., Claude, ChatGPT) and model version
- A summary table documenting total tokens consumed (input and output) per tool and per phase

The final term paper shall not exceed 20 pages (including R code and graphics) per group. **Groups usually consist of two students**. The **guidelines for scientific papers** at the Institute of Economics must be strictly adhered to. Formal aspects are part of the evaluation of the term paper. Students have to complete their seminar presentation and paper **in English**.

Seminar Instruction

Dr. Arne Steinkraus (Theory)/Sonja Kurzbach (Application and Organization)
Institut für Volkswirtschaftslehre
Spielmannstr. 9
38106 Braunschweig
a.steinkraus@tu-braunschweig.de/sonja.kurzbach@tu-braunschweig.de