Master's Seminar

Statistical Learning in Economics

Institut für Volkswirtschaftslehre
Winter Semester 2024-2025

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 of drugs. 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.

Prerequisites

Basic knowledge of Statistics and/or Econometrics is expected. Prior knowledge of Macroeconomics and Microeconomics is helpful. Familiarity with statistical programming languages such as R and Python is welcome but not required.

Requirements and Course plan

The following are expected from seminar participants:

- Attendance at the kick-off meeting

- Attendance in the R-lab exercise sessions

- Seminar presentation (15-minute presentation and 5-minute discussion)

- Active participation and contribution to other presentations

- Submission of the term paper

The seminar consists of four parts:

1. Kick-off and Theory crash-course: The session covers the basic concepts and theories of Statistical Learning. After the session, students can decide on a binding registration.

2. Application R-labs: There will be 03 R-lab sessions in which students are introduced to R and work on data analysis exercises to apply the theories covered in part 1.

3. Seminar course work (own project): For the seminar paper, students will conduct their own Statistical Learning analysis and write a term paper.

- The goal of the analysis is to predict the inflation rate in Germany for February 2025. To that end, students are expected to select, explain, train, and apply a suitable model.

- Additionally and critically, the tasks also include gathering own data and finding the relevant literature to substantiate the selected variables and give a theoretical introduction into the topic "Inflation". A list of suggested literature will be provided to students during the course as the starting point.

- The final term paper shall not exceed 20 pages (including R code and graphics) per group. Groups usually consist of two students. Grading will be based on the term paper only.

- The guidelines for scientific papers at the Institute of Economics can be found **here**. The formal guidelines must be strictly adhered to. Formal aspects are part of the evaluation of the term paper.

4. Block seminar: At the end of the course, students are required to give a short presentation of their term papers during the seminar.

Timeline

- Registration via Stud.IP: Open on 16.09.2024
- Kick-off and Theory crash-course: 18.10.2024 (13:00-17:00, Classroom to be confirmed)
- Application R-Labs: 25.10.2024, 01.11.2024, 08.11.2024 (9:00-12:00, VWL institute)
- Q&A Session: 22.11.2024 (10:00-12:00, Online)
- Submission deadline for the term paper: 10.01.2025 (16:00, via E-mail)
- Block seminar (Presentation): 17.01.2025 (9:00-12:00, VWL institute)

Content topics and literature

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

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
- Boostrapping

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

Literature

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). The elements of statistical learning: data mining, inference, and prediction (Vol. 2, pp. 1-758). New York: springer.

Giannone, D., Lenza, M., Momferatou, D., & Onorante, L. (2014). Short-term inflation projections: A Bayesian vector autoregressive approach. International journal of forecasting, 30(3), 635-644.

Kilian, L., & Zhou, X. (2022). The impact of rising oil prices on US inflation and inflation expectations in 2020–23. Energy Economics, 113, 106228.

Mankiw, G. (2017). Principles of Economics. (Recap on Macroeconomics)

Scheufele, R. (2011). Are qualitative inflation expectations useful to predict inflation?. OECD Journal: Journal of Business Cycle Measurement and Analysis, 2011(1), 29-53.

Smith, G. W. (2008). Japan's Phillips curve looks like Japan. Journal of Money, Credit and Banking, 40(6), 1325-1326.

Stock, J. H., & Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. Journal of economic literature, 41(3), 788-829.

Verbrugge, R. J., & Zaman, S. (2021). Whose inflation expectations best predict inflation?. Economic Commentary, (2021-19).

Seminar Instruction

Dr. Arne Steinkraus (Theory)
Institut für Volkswirtschaftslehre
Spielmannstr. 9
38106 Braunschweig
a.steinkraus@tu-braunschweig.de

Hanh My Le (Application and Seminar Organization)
Institut für Volkswirtschaftslehre
Spielmannstr. 9
38106 Braunschweig
hanh-my.le@tu-braunschweig.de