Unit 0

Overview

Objectives

After course completion students

- know about the statistical principles and characteristics of supervised learning techniques;
- are able to suitably apply supervised learning techniques to data sets.

Topics

- Regularized regression (lasso and elastic net).
- Model assessment and selection (cross-validation and bootstrap).
- Regression and classification trees.
- Ensemble methods (bagging, random forests, boosting).
- Deep learning with neural networks.

Organization

Prerequisites

- Computing
- Statistics I and II.
- Econometrics.

General information

- **Canvas WU**:
  Slides, assignments, course project.
- **Email:** Bettina.Gruen@wu.ac.at; Robert.Bajons@wu.ac.at
- **Literature:**
  - Trevor Hastie, Jerome H. Friedman, Robert Tibshirani (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition.
    PDF available for download at `https://web.stanford.edu/~hastie/ElemStatLearn/`.
  - Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani (2017) An Introduction to Statistical Learning: With Applications in R. Eigth Printing.
    PDF available for download at `https://www.statlearning.com/`.

Details

- Course dates
  - Lecture: Weekly, Tuesday morning.
  - Exercises: Weekly, Tuesday afternoon.
  - Project presentations: Tuesday / Thursday.
- Evaluation
  - Exercises.
  - Course project.
  - Written final exam.

Performance assessment

- Exercises (40%)
- Course project (20%)
- Written final exam (40%)
- Grading: after rounding

Percent	Grade
90 – 100	Very good
80 – 89	Good
65 – 79	Satisfactory
50 – 64	Adequate
0 – 49	Unsatisfactory

Lecture

Unit	Date	Time	Topic
1	2025-10-07	08:00–12:00	Introduction and overview; regularized regression (lasso and elastic net)
2	2025-10-14	08:00–11:00	Regularized generalized linear models
3	2025-10-21	08:00–11:00	Model assessment and selection (cross-validation and bootstrap)
4	2025-10-28	08:00–11:00	Regression and classification trees; ensemble methods: bagging, random forests
5	2025-11-04	08:00–11:00	Ensemble methods: boosting; deep learning with neural networks

Exercises

- 5 exercise units with 8 exercises per unit.
- Students prepare assignments at home and indicate by completing an assignment on Canvas WU with common deadline one hour before the exercise classes start which assignments they prepared and are able to present in class.
- During the exercise classes, students who ticked that they prepared the assignment via Canvas WU are randomly selected to present their solution of the assignment.
- For the presentation to be deemed suitable, the student must have a complete solution proposal where all points of the assignment are fully addressed. Incomplete solution proposals and insufficient presentation performance lead to all exercises ticked for this exercise class being canceled for this student.

Exercises / 2

- Evaluation
  - Best 4 out of 5 exercise units according to the number of exercises ticked count towards the grade (80%).
  - Presentation performance (20%).

Course project

- Team work in groups of 2 students (in case of an odd number: one group with 3 students). Please self-create groups in Canvas WU until 2025-10-13 23:59.
- Upload slides to Canvas WU on the day before the presentation.
- Oral presentation during the respective course unit.