Infra Informatics - Statistical Research Methods and Data Analysis, 2.0 credits

Infrainformatik - Statistiska metoder och dataanalys, 2.0 hp

Third-cycle education course

6FITN64

Department of Science and Technology

Valid from: First half-year 2025

Approved by
The Board of PhD Studies

Approved
2025-03-26

Registration number

Entry requirements

Admitted as a doctoral student.

Specific information

The course aims to

- give an overview of different statistical learning methods
- discuss methodology, tools, and practices for applied research in statistical learning
- present data-driven research projects within Infra Informatics.

Learning outcomes

After completing the course, participants should be able to:

- describe and categorize various statistical learning methods,
- describe and discuss the suitability, possibilities, and limitations of different statistical methods in relation to certain problem settings,
- identify, select, and plan the necessary steps for conducting a successful data-driven research project,
- describe and discuss data-driven research as well as the use of different statistical methods within Infra Informatics.

Contents

The following topics will be covered (in varying depth):

- Research methodology in data-driven research projects
- Overview of statistical methods and machine learning
- Estimation, prediction and inference
- Overfitting and bias-variance trade-off
- Supervised learning: different methods for regression and classification
- Unsupervised learning: clustering, dimensionality reduction and density estimation
- Resampling and evaluation methods
- Hypothesis testing
- Tools and programming languages for statistical data processing

Educational methods

The course is held during VT1 each year. The schedule consists of

- A half-day startup meeting, presenting the common framework for the whole course and first lectures on the course topics.
- Two full day seminar days, with lectures on the course topics, a lab session and presentations by senior researchers at KTS about projects where statistical methods have been applied.
- An individual homework assignment.
- A half-day final meeting, with student presentations and discussions.

Examination

Responsible for the Statistical Research Methods and Data Analysis course is David Gundlegård.
The examination for the course consists of:

- Mandatory participation in the four seminar occasions
- Conducting, documenting, and presenting an individual assignment

Examiner for the course is Mats Janné.

Grading

Two-grade scale

General information

The course is mandatory for all doctoral students in Infra Informatics. It is also open to doctoral students in other fields.

To register for the course, contact the course coordinator and Mats Janné.