SF2935 Modern Methods of Statistical Learning 7.5 credits

Moderna metoder för statistisk inlärning

This is a translation of the Swedish, legally binding, course syllabus.

Establishment

Course syllabus for SF2935 valid from Autumn 2017

Grading scale

A, B, C, D, E, FX, F

Education cycle

Second cycle

Main field of study

Mathematics

Specific prerequisites

Courses in probability and statistics, liner algebra, calculus in one and several variables, numerical methods.

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

This course presents an overview of the most important methods of the modern theory of statistical learning. This course focuses primarily on the practical aspects of statistical learning.Computer-aided project work with a variety of datasets forms the essential learning activity. To pass the course, the student should be able to do the following:

- explain the difference between unsupervised and supervised learning
- know the underlying mathematical relationships within and a cross statistical learning algorithms and the paradigms of supervised and unsupervised learning along with their strengths and weaknesses
- identify the correct statistical tool for a data analysis problem in the real world based on reasoned argument
- use algorithmic models treating the data mechanism as unknown
- develop accurate and informative alternatives to data modelling on big and complex as well as on smaller data sets
- design and implement various statistical learning algorithms in a range of real-world applications
- design test procedures in order to evaluate a model, optimise the models learned and report on the expected accuracy that can be achieved by applying the models
- read current research papers and understand the issues raised by current research.To receive the highest grade, the student should in addition be able to do the following:
- combine several models in order to gain better results.

Course contents

This course presents an overview of the most important methods of the modern theory of statistical learning. Topics covered include supervised learning with a focus on classification methods, support vector machines, artificial neural networks, decision trees, boosting, bagging and methods of unsupervised learning with focus on K-means clustering and nearest neighbours. This course focuses primarily on the practical aspects of statistical learning. Computer-aided project work with a variety of datasets forms the essential learning activity.

Disposition

Lectures, presentations, work with computer-aided data analysis.

Course literature

An introduction to Statistical Learning, by G. James, D. Witten, T. Hastie, R. Tibshirani, Springer Verlag, and additional reading available on the course web page.

Examination

- ÖVN1 - Assignments, 3.0 credits, grading scale: P, F
- TENA - Examination, 4.5 credits, grading scale: A, B, C, D, E, FX, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

If the course is discontinued, students may request to be examined during the following two academic years.

The written exam deals with concepts.

Other requirements for final grade

Written exam, assignments.

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.