SF1935 Probability Theory and Statistics with Application to Machine Learning 7.5 credits

Sannolikhetsteori och statistik med tillämpning inom maskininlärning

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

Establishment

The Head of School at the SCI school has 2021-10-15 decided to establish this syllabus to apply from Spring 2023, registration number: S-2021-1634

Grading scale

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

Education cycle

First cycle

Main field of study

Specific prerequisites

Completed course SF1625 Calculus in One Variable

Language of instruction

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

Intended learning outcomes

To pass the course, the student should be able to:

- solve problems that require knowledge about standard concepts and methods in probability theory
- solve problems that require knowledge about standard concepts and methods in statistical theory
- carry out a project work in a group with oral or written presentation and apply machine learning methods for data analysis problems

Course contents

Basic concepts such as probability, conditional probability and independent events. Discrete and continuous stochastic variables, especially one-dimensional stochastic variables. Location, spread and dependency measures for stochastic variables and data sets. Common distributions and their model situations, including the normal distribution, the binomial distribution and the Poisson distribution. The Central limit value theorem and the Law of large numbers.

Descriptive statistics. Point estimates and general estimation methods such as the Maximum likelihood method and the Minimum square method. General confidence intervals but special confidence intervals for expected value and variance in normal distribution. Confidence interval for participations and difference in expected values and participations. Hypothesis testing. Chi2 test of distribution, homogeneity test and independence test. Linear regression.

Machine learning paradigms, appoaches and applications. Supervised / unsupervised learning, generalization, model selection, validation and evaluation, probabilistic methods, dimensionality reduction and representations.

Examination

- TEN1 - Written exam, 6.0 credits, grading scale: A, B, C, D, E, FX, F
- PRO1 - Project work, 1.5 credits, grading scale: P, 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.

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.