Introduction to Machine Learning

Single subject and programme course

9 credits

Introduktion till maskininlärning

732A95

Valid from: 2016 Autumn semester

Determined by
The Quality Board at the Faculty of Arts and Sciences

Date determined
2016-04-01

Main field of study
Statistics

Course level
Second cycle

Advancement level
A1X

Course offered for
- Master´s Programme in Statistics and Data Mining

Entry requirements
A bachelor's degree in one of the following subjects: statistics, mathematics, applied mathematics, computer science, engineering, or equivalent. Completed courses in calculus and linear algebra are required.
The student should also have passed:
- a basic course in statistics covering at least 6 ECTS credits;
- a course in programming covering at least 6 ECTS credits.
Documented knowledge of English equivalent to Engelska B/Engelska 6.

Intended learning outcomes
After completion of the course the student should at an advanced level be able to:
- use relevant concepts and methods from machine learning in order to formulate, structure and solve practical problems that involve large or complex data,
- make an inference for the parameter values for commonly used machine learning models,
- use machine learning models for prediction and decision making,
- estimate the quality of the machine learning models,
- select a suitable model in situations with a limited or no information about the underlying dependencies in the data,
- implement machine learning models in a programming language and use existing machine learning software in order to analyze large and/or complex datasets, make predictions and estimate the uncertainty of these predictions.

Course content

The course introduces main concepts and tools in probabilistic machine learning which are necessary for professional work and research in data analytics.
- introduction to and overview of machine learning (including regression, classification, supervised and unsupervised learning) and its application areas,
- Nearest Neighbors and Naïve Bayes,
- discriminant analysis, logistic regression and decision trees,
- model selection and uncertainty estimation: holdout method, cross-validation, AIC, bootstrap confidence intervals,
- linear regression and regularization methods (Ridge, LASSO),
- splines, generalized linear and additive models,
- Principal component analysis (PCA) and Principal component regression (PCR),
- kernel smoothers, kernel trick and support vector machines,
- neural networks,
- bagging, boosting and random forests,
- Online learning and mixture models. .

Teaching and working methods

The teaching comprises lectures, seminars, and computer exercises, complemented by self-studies. Lectures are devoted to presentations of theories, concepts and methods. Computer exercises provide practical experience of data analysis in some machine learning software. The seminars comprise student presentations and discussions of computer assignments.
Language of instruction: English. .

Examination

Written reports on the computer assignments. Active participaton in the seminars. One final written examination. Detailed information about the examination can be found in the course's study guide.

Students failing an exam covering either the entire course or part of the course two times are entitled to have a new examiner appointed for the reexamination.

Students who have passed an examination may not retake it in order to improve their grades.

Students failing an exam covering either the entire course or part of the course twice are entitled to have a new examiner appointed for the reexamination.

Students who have passed an examination may not retake it in order to improve their grades.

Grades

Three-grade scale, U, G, VG

Other information

Planning and implementation of a course must take its starting point in the wording of the syllabus. The course evaluation included in each course must therefore take up the question how well the course agrees with the syllabus.

The course is carried out in such a way that both men´s and women´s experience and knowledge is made visible and developed.

Planning and implementation of a course must take its starting point in the wording of the syllabus. The course evaluation included in each course must therefore take up the question how well the course agrees with the syllabus.

The course is carried out in such a way that both men´s and women´s experience and knowledge is made visible and developed.

Department

Institutionen för datavetenskap