CM2007 Applied Machine Learning and Data Mining for Performance Analysis 7.5 credits

Tillämpad maskininlärning och datautvinning för prestationsanalys

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

Establishment

Course syllabus for CM2007 valid from Spring 2024

Grading scale

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

Education cycle

Second cycle

Main field of study

Technology and Health

Specific prerequisites

At least 10 credits in linear algebra and analysis, at least 6 credits in mathematical statistics, and at least 8 credits in object-oriented programming.

Language of instruction

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

Intended learning outcomes

On completion of the course, the student is expected to be able to:

•apply methods to import, combine and convert data into the appropriate format for data analysis,

•explain the benefits of data mining and be able to choose and implement appropriate methods in typical data mining use cases,

•choose, motivate, and apply standard machine learning methods and algorithms to typical use cases and present the results in appropriate ways

•design and perform performance validation of a machine learning system

•give an account on ethics and regulations when using and processing personal data.

Course contents

This course deals with how to process and draw conclusions of data through data mining and machine learning. The course introduces some theory on machine learning, but focuses mainly on current applied methods.

The following is included in the course:

•Statistical and probabilistic methods for data analysis.

•Different methods for data mining.

•Algorithms for supervised and unsupervised machine learning.

•Neural networks and deep learning.

•Data extraction: purpose and typical use cases in performance analysis.

•Routines for importing, combining, converting and selecting data for learning and validation.

•Validation methods and performance measures.

•Visualization and analysis of results from data analysis.

•Ethics and regulations concerning use and processing of personal data.

Examination

- LAB1 - Laboratory work, 4.0 credits, grading scale: A, B, C, D, E, FX, F
- RED1 - Reporting, 2.0 credits, grading scale: P, F
- TEN1 - Oral examination, 1.5 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.

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.