



# DD1418 Language Engineering with Introduction to Machine Learning 6.0 credits

Språkteknologi med introduktion till maskininlärning

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

## Establishment

Course syllabus for DD1418 valid from Autumn 2019

## Grading scale

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

## Education cycle

First cycle

## Main field of study

Technology

## Specific prerequisites

Completed course in computer science equivalent to DD1320/DD1321 Applied computer science, DD1327 Fundamentals of computer science or DD1338 Algorithms and Data Structures.

Completed course in probability theory and statistics equivalent to SF1918.

# Language of instruction

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

## Intended learning outcomes

Having passed the course, the student should be able to:

1. explain and use basic concepts in linguistics, language engineering and machine learning
  2. apply language engineering concepts, methods and tools to build language engineering systems as well as be able to explain the structure of such systems
  3. implement standard methods in language engineering
  4. design and carry out simple evaluations of a language engineering system as well as interpret the results,
  5. independently be able to solve a well delimited practical language engineering problem in order to be able to
- work with a bachelor's degree project with a focus on language engineering or machine learning,
  - be an important link between systems designers, programmers, and interaction designers in industry as well as in research projects.

## Course contents

Theory:

The historical development and bases of language engineering, morphology, syntax, semantics, vector space models, evaluation methods, machine learning, information theory and Markov models.

Methods::

Morphological analysis, generation and language statistics and corpus processing, parsing, generation, part-of-speech tagging, named entity recognition, probabilistic parsing and statistical lexical semantics.

Application areas:

Spelling and grammar checking, information retrieval, word prediction for smart text entry, text clustering and text categorization, computer-aided language learning, dialogue systems, speech technology and machine translation.

## Disposition

Theoretical lectures and applied lectures interleaved with practical laboratory sessions. A final project work that is presented orally and in writing.

## Course literature

The reading list is announced on the course page.

## Examination

- PRO1 - Project, 1.5 credits, grading scale: A, B, C, D, E, FX, F
- TEN1 - Exam, 3.0 credits, grading scale: A, B, C, D, E, FX, F
- LAB1 - Laboratory Assignments, 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.

It is the examiner who, in agreement with KTH's coordinator for disabilities, will decide on possible adapted examination for students with a proven permanent disability. The examiner may permit another examination format at the re-examination of individual students.

## Other requirements for final grade

Passed laboratory course, project assignment and exam.

## 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.