# 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.

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

## Establishment

Course syllabus for DD1418 valid from Autumn 2017

## Grading scale

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

## Education cycle

First cycle

## Main field of study

Technology

## Specific prerequisites

A course in computer science equivalent to DD1320/DD1321 Applied computer science, DD1327 Fundamentals of computer science or DD1338 Algorithms and Data Structures. A course in probability theory and statistics equivalent to SF1901.

## Language of instruction

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

## Intended learning outcomes

After a pass mark on course, the students should be able to:

â€¢ explain and use concepts in the following levels of linguistics: morphology, syntax, semantics, discourse and pragmatics,

â€¢ apply knowledge of morphology, syntax and lexical semantics to build language engineering systems as well as explain the structure of existing system based on these levels

â€¢ use basic tools in language engineering such as part-of-speech taggers, chunkers as well as various types of corpora and dictionary to be able to build own programs,

â€¢ explain and use standard methods in language technology that are based on both rules, statistics and machine learning,

â€¢ apply methods based on finite automata/transducers, context-free grammars, word frequencies, n-grams, co-occurrence statistics, Markov models, and vector space models,

â€¢ analyze and explain which problems within language technology that could be solved with usable results, and which could not be solved,

â€¢ explain how spelling and grammar checkers, taggers based on machine learning, stemmers, and an algorithm for semantic content acquisition work,

â€¢ design and carry out simple evaluations of some language engineering system as well as interpret the results

â€¢ independently solve a well-defined practical language technology problem, or analyze a problem theoretically,

in order to:

â€¢ work with a bachelor's degree project with a focus on language engineering or machine learning,

â€¢ 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.

Technologies:

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 fields:

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

Will be announced on the course web no later than 10 weeks before the start of the course.

## Examination

- LAB1 - Laboratory Assignments, 1.5 credits, grading scale: P, F
- 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

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.

Under special circumstances, other examination formats may be used.

In this course, the code of honor of the school is applied, see: http://www.kth.se/en/csc/utbildning/hederskodex

## Other requirements for final grade

Passed laboratory course, project assignment and examination.

## 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.