



DT2119 Speech and Speaker Recognition 7.5 credits

Igenkänning av tal och talare

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

On 2019-10-15, the Head of School of EECS has decided to establish this official course syllabus to apply from the spring semester 2020 (registration number J-2019-0336).

Grading scale

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

Education cycle

Second cycle

Main field of study

Computer Science and Engineering

Specific prerequisites

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 shall be able to

- implement methods for training and evaluation of speech recognition systems
- train and evaluate a speech recognizer, using software tools
- compare different methods for feature extraction and training
- document and discuss specific aspects related to recognition of speech and of speakers
- review and criticise other students' work in the subject, based on the literature.

Course contents

The course consists of lectures, three laboratory sessions with hand-in assignments, as well as writing an essay on a subject chosen in consultation with the teacher. The thesis is furthermore presented orally during a final seminar. The laboratory sessions consist of designing different parts of a speech recognition application, training the system and evaluating its performance.

The following theoretical course components are included:

- algorithms for training, recognition as well as adaptation to properties of speakers and transmissions channel, including pattern recognition, Hidden Markov Models (HMMs) and Deep Neural Networks (DNNs)
- methods to decrease the sensitivity to disturbances and deviations
- probability theory
- signal processing and parameter extraction
- acoustic modelling of the static and dynamic spectral properties of speech sounds
- statistical modelling of language in spontaneous and formal speech
- search strategies - basic methods and strategies for large vocabularies
- specific methods for analysis and decision making, for recognition of speakers.

Furthermore, some practical insights into building an application are given. This includes the implementation of certain functions based on prototypes, and testing them on real speech data.

Examination

- LAB1 - Computer Lab, 4.5 credits, grading scale: P, F
- PRO1 - Project, 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.

Other requirements for final grade

Laboratory exercises

Written assignments.

Academic paper and its presentation at a final review

Assessment of two other course participants' theses, and critical review of their presentations.

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.