UNIVERSITY OF GOTHENBURG

DEPARTMENT OF PHILOSOPHY, LINGUISTICS AND THEORY OF SCIENCE

LT2114 Practical natural language processing, 7.5 credits

Praktisk grundläggande språkteknologi, 7,5 högskolepoäng

Second Cycle

Confirmation

This course syllabus was confirmed by Department of Philosophy, Linguistics and Theory of Science on 2017-06-01 and was last revised on 2019-03-14 to be valid from 2019-03-18, spring semester of 2019.

Field of education: Science 100%
Department: Department of Philosophy, Linguistics and Theory of Science

Position in the educational system

Can also be offered as a freestanding course.

The course can be part of the following programmes: 1) Computer Science, Master's Programme (N2COS) and 2) Applied Data Science Master's Programme (N2ADS)

Main field of studies	*Specialization*
Language Technology	A1N, Second cycle, has only first-cycle course/s as entry requirements

Entry requirements

For admission to course is required at least 7.5 credits in programming such as DIT948 Programming, DIT012 Imperative programming with basic object orientation, DIT142 Functional programming, LT2001 Introduction to programming, or the equivalent.

Learning outcomes

On successful completion of the course the student will be able to:

Knowledge and understanding

- account for basic concepts in language technology,
- explain in detail some common algorithms for language technology assignments such as text categorisation or part-of-speech tagging,
- account for different evaluation methods in language technology as well as when they are applicable.

Competence and skills

- use a language technology software library to develop practical data driven language technology applications,
- implement methods for basic text processing,
- implement evaluation procedures for language technology assignments.

Judgement and approach

- discuss which language technology method is applicable to solve a specific language technology problem,
- make an informed assessment on how to solve an NLP task.

Course content

The course gives a practical introduction to different problems encountered within natural language processing and some standard solutions. Students will gain practical experience in programming while solving these problems. The course is divided into four main topics, one covering basic concepts and three covering subfields of natural language processing: basic text processing, syntactic and morphological analysis, as well as semantics and pragmatics. The final selection of the subjects covered in the four subfields is decided by the course coordinator before the start of the course. The following is an example of subjects that can be completed during the course:

1. Basic concepts:
   - basic concepts in language technology,
   - probability theory and machine learning for language technology problems,
   - common evaluation measurement, e.g. correctness, precision and recall.
2. Basic text processing:
   - corpora and corpus annotation,
   - basic statistical corpus analysis.
3. Syntax and morphology:
   - morphological analysis (part-of-speech analysis) with machine learning and finite-state technology,
   - syntactic analysis by means of rule-based and data-driven methods.
4. semantics and pragmatics:

- word sense disambiguation with machine learning,
- text classification or topic modeling with machine learning.

Sub-courses

1. Written examination (Written examination), 4 credits

Grading scale: Pass with distinction (VG), Pass (G) and Fail (U)

2. Assignments, 3.5 credits

Grading scale: Pass with distinction (VG), Pass (G) and Fail (U)

Form of teaching

Lectures, exercises, computer exercises. There are mandatory laboratory exercises that require attendance for a passing grade.

Language of instruction: English

Assessment

The course is assessed with a combination of written examination, written assignments, lab sessions and active participation while making exercises and attending laboratory sessions.

When a student, who has failed the same examined component twice, wishes to change examiner before the next examination, a written application shall be sent to the department responsible for the course and shall be granted unless there are special reasons to the contrary (Chapter 6, Section 22 of Higher Education Ordinance).

In the case where a course has been discontinued or has undergone major changes, the student shall normally be guaranteed at least three examination sessions (including the regular examination session) during a period of at least one year on the basis of the course's former structure.

Grades

The grading scale comprises: Pass with Distinction (VG), Pass (G) and Fail (U).
For the grade Pass is required:

- passed written examination
- completed written assignments
- active attendance at labs

To pass with distinction is required:

- written examination passed with distinction

- completed written assignments of high quality
- active attendance at labs

Course evaluation

The director of studies or the equivalent is responsible to ensure that a short version of the students course evaluation is done, published and taken into account when developing/improving the course.

The result and any changes in the course structure must be communicated to both the students who completed the evaluation and to the students who will start the course.