Text Mining

Text Mining
6 credits

Single subject and programme course

732A81

Valid from: 2025 Autumn semester

Determined by	**Main field of study**	
Course and Programme Syllabus Board at the Faculty of Arts and Sciences	Statistics	
Date determined	**Course level**	**Progressive specialisation**
2022-03-07	Second cycle	A1F
Revised by	**Disciplinary domain**	
Chairman of the Course and Programme Syllabus Board at the Faculty of Arts and Sciences	Technology	
Revision date	**Subject group**	
2024-09-30	Statistics	
Offered first time	**Offered for the last time**	
Autumn semester 2022		
Department	**Replaced by**	
Institutionen för datavetenskap		

Course offered for

- Master's Programme in Statistics and Machine Learning
- Master Programme in Cognitive Science

Entry requirements

- 180 ECTS credits passed including 90 ECTS credits in one of the following subjects:
    - statistics
    - mathematics
    - applied mathematics
    - computer science
    - engineering
- Passed courses in:
    - calculus
    - linear algebra
    - statistics
    - programming
- English corresponding to the level of English in Swedish upper secondary education (Engelska 6)
  Exemption from Swedish
- At least 24 ECTS credits passed in the main field of Statistics at second cycle and at least 5 ECTS credits passed in the main field of Computer Science at second cycle

Intended learning outcomes

After completing the course, the student shall be able to

- implement text mining methods and apply them to practical problems
- analyze and summarize results from text mining experiments
- identify, formulate and solve problems within the area of text mining
- clearly present and discuss the conclusions of a project work

Course content

The course covers the following methods within the area of text mining:

- information retrieval
- basic methods in language technology
- predictive modeling, in particular text classification
- text clustering and theme modeling
- information extraction
- validation methods

In addition, various applications of text mining as well as program libraries and data used in the field are treated.

Teaching and working methods

The teaching consists of lectures, computer labs and supervision in connection with a project work. In addition, the student shall engage in self-studies.

Language of examination: English

The student is entitled to supervision during the study period to complete the project work for the current course that the student is registered to.

Examination

The course is examined by:

- written and oral presentation of computer labs in groups, grading scale: UG
- individually written project report, grading scale: EC

For Pass (E) as a final grade, Pass (G) is required on the computer labs and at least E on the individually written project report. A higher grade is based on the individually written project report.

If special circumstances prevail, and if it is possible with consideration of the nature of the compulsory component, the examiner may decide to replace the compulsory component with another equivalent component.

If the LiU coordinator for students with disabilities has granted a student the right to an adapted examination for a written examination in an examination hall, the student has the right to it.

If the coordinator has recommended for the student an adapted examination or alternative form of examination, the examiner may grant this if the examiner assesses that it is possible, based on consideration of the course objectives.

An examiner may also decide that an adapted examination or alternative form of examination if the examiner assessed that special circumstances prevail, and the examiner assesses that it is possible while maintaining the objectives of the course.

Students failing an exam covering either the entire course or part of the course twice are entitled to have a new examiner appointed for the reexamination.

Students who have passed an examination may not retake it in order to improve their grades.

Grades

ECTS, EC

Other information

Planning and implementation of a course must take its starting point in the wording of the syllabus. The course evaluation included in each course must therefore take up the question how well the course agrees with the syllabus.

The course is conducted in such a way that there are equal opportunities with regard to sex, transgender identity or expression, ethnicity, religion or other belief, disability, sexual orientation and age.

If special circumstances prevail, the vice-chancellor may in a special decision specify the preconditions for temporary deviations from this course syllabus, and delegate the right to take such decisions.

About teaching and examination language

The teaching language is presented in the Overview tab for each course. The examination language relates to the teaching language as follows:

- If teaching language is “Swedish”, the course as a whole could be given in Swedish, or partly, or as a whole, in English. Examination language is Swedish, but parts of the examination can be in English.
- If teaching language is “English”, the course as a whole is taught in English. Examination language is English.
- If teaching language is “Swedish/English”, the course as a whole will be taught in English if students without prior knowledge of the Swedish language participate. Examination language is Swedish or English depending on teaching language.