Syllabus

Master’s Programme in Statistics and Data Mining
120 ECTS Credits

Aim
The rapid growth of databases provides scientists and business people with vast new resources. This programme meets the challenges of turning large or complex data sets into knowledge. Statistical modelling and analysis is integrated with machine learning, data mining and visualization into a solid basis for professional work with the organization and analysis of data, or a career in research.

Learning outcomes
Upon completing the programme the students shall be able to:
• extract and organize large volumes of complexly structured data
• explore, summarize and present large and complex data sets by static, interactive and dynamic graphical facilities
• select a suitable model for a given statistical problem and dataset
• uncover and statistically verify previously unknown patterns and trends in the data
• use advanced statistical and data mining computer software to analyse large data volumes
• implement models suitable for data analysis in some computer language
• combine data information with other sources of prior information to improve inference and prediction performance
• give examples of application areas where analysis of large and complex data sets is needed
• present a written thesis with a theoretical or an applied study of a complex data set

Knowledge and understanding
Upon completing the programme the student shall
• demonstrate knowledge and understanding in statistics, including both broad knowledge of the field and a considerable degree of specialised knowledge in its branch, data mining, as well as insight into current research and development work, and
• demonstrate specialised methodological knowledge in statistics.

Specialized knowledge in data mining shall include modern powerful techniques for prediction, classification, clustering, Bayesian methods and association analysis.

Competence and skills
Upon completing the programme the student shall
• demonstrate the ability to critically and systematically integrate knowledge and analyse, assess and deal with complex phenomena, issues and situations even with limited information
• demonstrate the ability to identify and formulate issues critically, autonomously and creatively as well as to plan and, using appropriate methods, undertake advanced tasks
within predetermined time frames and so contribute to the formation of knowledge as well as the ability to evaluate this work

- demonstrate the ability in speech and writing both nationally and internationally to report clearly and discuss his or her conclusions and the knowledge and arguments on which they are based in dialogue with different audiences, and
- demonstrate the skills required for participation in research and development work or autonomous employment in some other qualified capacity.

**Judgement and approach**

Upon completing the programme the student shall

- demonstrate the ability to make assessments in statistics informed by relevant disciplinary, social and ethical issues and also to demonstrate awareness of ethical aspects of research and development work
- demonstrate insight into the possibilities and limitations of research, and especially research in statistics and data mining, its role in society and the responsibility of the individual for how it is used, and
- demonstrate the ability to identify the personal need for further knowledge and take responsibility for his or her ongoing learning.

**Content**

The curriculum joins courses in statistics, computer science and mathematics into a programme in the interface between statistics and computer science. Compulsory courses, introductory courses, and a 30-credit master’s thesis ensure progression and depth. Introductory courses are offered to fill in knowledge gaps and ensure that the students are properly prepared for the other courses.

**Compulsory Courses**

**Advanced Academic studies, 3 credits (given in semester 1)**
The aim of the course is to prepare the students for advanced academic studies and also to let the students learn the academic culture in general. A basic ambition is to supply essential tools to the students on the master’s level in Sweden. In addition, practical issues that are specific for the programme are to be discussed.

**Introduction to Machine Learning, 9 credits (given in semester 1)**

**Advanced Data Mining, 6 credits (given in semester 2)**

**Big Data Analytics, 6 credits (given in semester 2)**
**Introduction to Python, 3 credits (given in semester 2)**

**Philosophy of Science, 3 credits (given in semester 2)**
Laws of nature and scientific models. Relations between theories and observations. Forces prompting scientific change.

**Bayesian Learning, 6 credits (given in semester 2)**
Bayes' theorem to combine data information with other prior information. Bayesian analysis of conjugate models. Markov Chain Monte Carlo methods for Bayesian computations. Bayesian model comparison.

**Computational statistics, 6 credits (given in semester 2)**

### Profile courses

**Visualization, 6 credits (given in semester 1 for students admitted in an even year and in semester 3 admitted in an odd year)**
Advanced visualization techniques for large and complex data sets. Interactive and dynamic statistical graphics. Visualizing spatial information.

**Advanced Machine Learning, 6 credits (given in semester 3)**

**Time Series Analysis, 6 credits (given in semester 1 for students admitted in an odd year and in semester 3 admitted in an even year)**

**Multivariate Statistical Methods, 6 credits (given in semester 1)**
Analysis of correlation and covariance structures, including principal components, factor analysis and canonical correlation. Classification and discrimination techniques. Multivariate inference.

**Probability Theory, 6 credits (given in semester 3)**

**Decision Theory, 6 credits (given in semester 3)**

### Complementary courses
Web Programming, 6 credits (given in semester 2)
Overview of WWW, HTML, Javascript and other client-server techniques. Programming languages Python, Flask, SQL, Websockets, JSON and other server-side technologies

Neural networks and learning systems, 6 credits (given in semester 2)

Data Mining Project, 6 credits (given in semester 3)
Project course in which the student specifies, implements and evaluates a data mining algorithm for a specific data mining problem.

Text mining, 6 credits (given in semester 3)
Retrieval of textual data from different sources. Text processing by means of computational linguistics. Statistical models for text classification and prediction.

Database Technology, 6 credits (given in semester 3)

Introductory courses

Statistical methods, 6 credits (given in semester 1)

Advanced R programming, 6 credits (given in semester 1)

Master’s thesis, 30 credits
Theoretical or applied study of a complex data set by using statistical, machine learning and data mining methods.

Admission Requirements

General requirements
A person meets the general entry requirements for courses or study programmes that lead to the award of a second-cycle qualification if he or she:

1. possesses a first-cycle qualification comprising at least 180 credits or a corresponding qualification from abroad, or
2. by virtue of courses and study programmes in Sweden or abroad, practical experience or some other circumstance has the aptitude to benefit from the course or study programme.

**Specific requirements**

**Knowledge of English**
Documented knowledge of English equivalent to "Engelska B"; i.e. English as native language or an internationally recognized test, e.g. TOEFL (minimum scores: Paper based 575 + TWE-score 4.5, and internet based 90), IELTS, academic (minimum score: Overall band 6.5 and no band under 5.5), or equivalent.

**Degree results**
The specific requirements will be assessed as not fulfilled if the average grade is in the lower third of the grading scale used in the country where the degree was awarded, that is grades have to be average/pass or above (the equivalent to the Swedish grade “Godkänd”)

**Letter of Intent**
Each applicant must enclose a Letter of Intent, written in English by the applicant, comprising a motivation why the applicant wishes to follow the programme, and a summary of degree thesis/degree project.
For those holding a degree that does not require such a degree thesis/degree project the Letter of Intent should describe previous studies and academic activities related to the Master’s programme/es applied for.

**Programme Specific requirements**
A bachelor’s degree in one of the following subjects: statistics, mathematics, applied mathematics, computer science, engineering or a similar degree.
Courses in calculus, linear algebra, statistics and programming are also required.

**Teaching Methods and Examination**

**Teaching Methods**
Ordinary courses have lectures, seminars, and computer exercises. The lectures are devoted to presentations of theories, concepts, and methods. The seminars comprise presentations and discussions of assignments. The computer exercises provide practical experience of data analysis and other methods taught in the programme. The courses that are named projects have supervision only.

**Examination**
Ordinary courses yielding a minimum of 4.5 credits have one or more assignments and one written examination. Project courses and the master’s thesis are examined through a written report and oral defence of that report.
Grades
As stipulated in the course syllabi.

Transfer of Credits
The Board of the Faculty of Arts and Sciences or person nominated by the Board decides whether or not previous education can be transferred into the programme.

Certificates
The student will be awarded the degree of Master of Science (120 ECTS credits) in Statistics provided all course requirements are completed and that the student fulfils the general and specific eligibility requirements including proof of holding a Bachelor’s (kandidat) or a corresponding degree.
To be awarded the degree the students must have passed 90 ECTS credits of courses including 42 ECTS credits of the compulsory courses, a minimum of 6 ECTS credits of the introductory courses, a minimum of 12 ECTS credits of the profile courses, and, possibly, some amount of complementary courses. The students must also have successfully defended a master’s thesis of 30 ECTS credits.
Completed courses and other requirements will be listed in the degree certificate.
A degree certificate is issued by the Board of the Faculty of Arts and Sciences on request.

Enrolment Procedure
Students are admitted to the programme in its entirety.

Regulations for semester admission
The student must have passed at least 6 ECTS credits of the first semester, in order to be admitted to the second semester of the programme.
The student must have passed at least 40 ECTS credits of the first year in order to be admitted to the third semester of the programme.
The student must have passed at least 65 ECTS credits of the programme, including all obligatory courses, in order to be admitted to the fourth semester of the programme.

Language of instruction
The language of instruction is English.

The Study Programme Syllabus was approved by the Board of the Faculty of Arts and Sciences on 2015, June 16 and changed September 2016.
The syllabus is valid from Autumn 2017.

LiU-2015-01422