

Big Data Analytics

Single subject and programme course

6 credits

Analys av Big data

732A54

Valid from: 2016 Autumn semester

Determined by
Board of the Faculty of Arts and Sciences

Date determined
2016-09-30

Main field of study

Statistics

Course level

Second cycle

Advancement level

A1X

Course offered for

- Masters Programme in Statistics and Machine Learning

Entry requirements

A bachelor's degree in one of the following subjects: statistics, mathematics, applied mathematics, computer science, engineering, or equivalent. Completed courses in calculus and linear algebra are required. Completed courses in statistics covering at least 6 ECTS credits and a course in programming covering at least 6 ECTS credits are also required.

Documented knowledge of English equivalent to Engelska B/Engelska 6.

Intended learning outcomes

After completed the course, the student should on an advanced level be able to:

- collect and store Big Data in a distributed computer environment
- perform basic queries to a database operating on a distributed file system
- account for basic principles of parallel computations
- use MapReduce concept to parallelize common data processing algorithms
- account for how standard machine learning models should be modified in order to process Big Data
- use tools for machine learning for Big Data

Course content

The course introduces main concepts and tools for storing, processing and analyzing Big Data which are necessary for professional work and research in data analytics.

- Introduction to Big Data: concepts and tools
- Introduction to Python
- Basic principles of parallel computing
- Introduction to databases
- File systems and databases for Big Data
- Querying for Big Data
- Resource management in a cluster environment
- Parallelizing computations for Big Data
- Basic Machine Learning algorithms
- Machine Learning for Big Data

Teaching and working methods

The teaching comprises lectures and computer exercises. Lectures are devoted to presentations of theories, concepts and methods. Computer exercises provide practical experience of manipulation with Big Data. Homework and independent study are a necessary complement to the course. Language of instruction: English.

Examination

Written reports on the computer assignments. Written examination. Detailed information about the examination can be found in the course's study guide.

Grades

ECTS, EC

Department

Institutionen för datavetenskap