Big Data Analytics

Programme course

6 credits

Big Data Analytics

TDDE31

Valid from: 2021 Spring semester

Determined by

Board of Studies for Computer Science and
Media Technology

Date determined

2020-09-29

Main field of study

Information Technology, Computer Science and Engineering, Computer Science

Course level

Second cycle

Advancement level

A1X

Course offered for

- Master's Programme in Computer Science
- Computer Science and Engineering, M Sc in Engineering
- Industrial Engineering and Management - International, M Sc in Engineering
- Industrial Engineering and Management, M Sc in Engineering
- Information Technology, M Sc in Engineering
- Computer Science and Software Engineering, M Sc in Engineering
- Applied Physics and Electrical Engineering - International, M Sc in Engineering
- Applied Physics and Electrical Engineering, M Sc in Engineering
- Master's Programme in Mathematics

Prerequisites

Basic database course. Data mining or machine learning course.

Intended learning outcomes

After completed 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 the MapReduce concept to parallelize common data processing algorithms
- be able to modify standard machine learning models 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
- Basic principles of parallel computing
- File systems and databases for Big Data
- Querying for Big Data
- Resource management in a cluster environment
- Parallelizing computations for Big Data
- 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.

Examination

LAB1	Labs	U, G	3 credits
TEN1	Written exam	U, 3, 4, 5	3 credits

Grades

Four-grade scale, LiU, U, 3, 4, 5

Course literature

Article collection.

Other information

Related courses: advanced data models and databases, parallel programming, multicore programming.

Department

Institutionen för datavetenskap

Director of Studies or equivalent

Patrick Lambrix

Examiner

Olaf Hartig

Course website and other links

http://www.ida.liu.se/~TDDE31/

Education components

Preliminär schemalagd tid: 42 h
Rekommenderad självstudietid: 118 h

Course literature

Övrigt

Artikelsamling 2018.

LINKÖPING UNIVERSITY