

Big Data Science (E018210)

Course size (nominal values; actual values may depend on programme)

Credits 6.0 Study time 180 h Contact hrs 60.0 h

Course offerings and teaching methods in academic year 2020-2021

A (semester 2)	English	lecture	30.0 h
		practicum	30.0 h

Lecturers in academic year 2020-2021

De Bie, Tijl	TW06	lecturer-in-charge
Mannens, Erik	TW06	co-lecturer

Offered in the following programmes in 2020-2021

	crdts	offering
Brugprogramma Master of Science in Bioinformatics (main subject Engineering)	6	A
Bridging Programme Master of Science in Industrial Engineering and Operations Research	6	A
Bridging Programme Master of Science in Industrial Engineering and Operations Research	6	A
Master of Science in Business Engineering (main subject Data Analytics)	6	A
Master of Science in Bioinformatics (main subject Engineering)	6	A
Master of Science in Business Engineering (main subject Operations Management)	6	A
Master of Science in Industrial Engineering and Operations Research	6	A
Master of Science in Computer Science Engineering	6	A
Master of Science in Computer Science Engineering	6	A
Master of Science in Industrial Engineering and Operations Research	6	A
Exchange Programme in Bioinformatics (master's level)	6	A
Exchange Programme in Computer Science (master's level)	6	A
Postgraduate in Internet of Things	6	A

Teaching languages

English

Keywords

NoSQL, Big Data platforms & architecture, Interactive data visualizations, Semantic Web technologies, large-scale data mining & machine learning, stream management, streaming algorithms, recommender systems

Position of the course

The main purpose of this course is to let the students gain hands-on experience with the most important concepts in Big Data Science. They will learn how to manage, analyze and visualize Big Data.

Contents

- Big Data Management:
- NoSQL stores, CAP Theorem
 - Platforms: Hadoop & Spark
 - Stream management
- Interactive data visualizations:
- Human perception

- Design principles & interaction
 - D3.js & Leaflet
- Algorithms & Analytics:
- Algorithms on the web (PageRank, Adwords,...)
 - Machine Learning (recommender systems, classification, regression,...)
 - Data Mining (dimensionality reduction, clustering,...)
 - Stream mining
- Guest Lectures from Belgian Big Data companies

Initial competences

The lab sessions and the project require basic python and javascript knowledge. The course can be taken by students who only have programming experience in other languages than those, but will then require some extra effort.

Final competences

- 1 Understand the components of (big) data systems and how they can be combined into a system architecture meeting the needs of a specific data product.
- 2 Being able to perform the full Big Data life cycle: cleansing & (visually) exploring data, creating a Big Data model, creating interactive dashboards.
- 3 A solid understanding of scalable data mining & machine learning algorithms.
- 4 Being able to handle datasets which are challenging along many dimensions: size, data format, a lack of data quality.

Conditions for credit contract

Access to this course unit via a credit contract is determined after successful competences assessment

Conditions for exam contract

This course unit cannot be taken via an exam contract

Teaching methods

Lecture, practicum

Extra information on the teaching methods

Lectures, practicums, project

Learning materials and price

PowerPoint presentations, articles and book chapters (freely available online)

References

- Mining of Massive Datasets, Jure Leskovec, Anand Rajaraman, Jeffrey Ullman, ISBN: 978-1-107-07723-2
- Big Data: A Revolution That Will Transform How We Live, Work, and Think, Viktor Mayer-Schönberger, ISBN 978-0-544-00269-2
- Predictive Analytics, Eric Siegel, ISBN 978-1-118-35685-2
- Doing Data Science, Rachell Schutt, ISBN 978-1-449-35865-5

Course content-related study coaching

Evaluation methods

end-of-term evaluation and continuous assessment

Examination methods in case of periodic evaluation during the first examination period

Oral examination

Examination methods in case of periodic evaluation during the second examination period

Oral examination

Examination methods in case of permanent evaluation

Skills test, report

Possibilities of retake in case of permanent evaluation

examination during the second examination period is possible in modified form

Extra information on the examination methods

- During examination period: oral closed-book exam
- During semester: graded lab session reports; graded project reports (approximately 8 lab/project reports in total)
- Second chance: possible in adapted form; Frequency: certainly graded lab session & 1 reading assignment

Calculation of the examination mark

- Evaluation throughout semester as well as during examination period.

- Special conditions: Global score = 50% score obtained for work during the semester (lab/project reports) + 50% score obtained for the end-of-semester exam.
- Additional requirement to pass: having obtained at least 8/20 for each of these parts (work during the semester & end-of-semester exam).