



COURSE SYLLABUS **Data Science, 7.5 credits**

Data Science, 7,5 högskolepoäng

Course Code: TDSR29	Education Cycle: Second-cycle level
Confirmed by: Dean Jun 1, 2019	Disciplinary domain: Technology
Valid From: Aug 1, 2019	Subject group: DT1
Version: 1	Specialised in: A1N
	Main field of study: Informatics

Intended Learning Outcomes (ILO)

After a successful course, the student shall

Knowledge and understanding

- Display knowledge of the fundamental tasks in data analysis, i.e., classification, regression, clustering and association rules
- Display knowledge of the phases in a data analysis project, i.e. preprocessing, modeling and evaluation
- Display knowledge of tools and techniques for large scale data analysis
- Show familiarity with key research directions in data analysis and machine learning

Skills and abilities

- Demonstrate the ability to use a software tool for all parts of a data analysis project; i.e., preprocessing, modeling and evaluation
- Demonstrate the ability to identify and apply a suitable technique based on a problem description
- Demonstrate the ability to analyse results and models from a data analysis project

Judgement and approach

- Demonstrate an understanding of how data analysis can be used as a tool to support business and technical decision-making

Contents

The exponential growth of the digital universe, particularly in the form of storage and computing power in recent decades, enables companies to accumulate huge amounts of data at moderate cost. Accompanying this technological shift is a widespread realization that collected data contain potentially valuable information. Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data in various forms, both structured and unstructured, through data analysis.

The course includes the following elements:

- Fundamental tasks in data analysis, i.e., classification, regression, clustering and association rules
- Basic machine learning techniques for classification, regression, clustering and association rules
- Organization of a data analysis project, i.e., preprocessing, modeling and evaluation
- Tools and techniques for large scale data analysis
- Software tools for data analytics
- Data analytics applied in different engineering and business domains

Type of instruction

The course consists of lectures, seminars and assignments with tutoring.

The teaching is conducted in English.

Prerequisites

The applicant must hold the minimum of a bachelor's degree (i.e the equivalent of 180 ECTS credits at an accredited university) with at least 90 credits in computer engineering, electrical engineering (with relevant courses in computer engineering), or equivalent. The bachelor's degree should comprise a minimum of 15 credits in mathematics. Proof of English proficiency is required.

Examination and grades

The course is graded 5,4,3 or Fail.

The final grade for the course is based on a balanced set of assessments. The final grade will only be issued after satisfactory completion of all assessments.

Registration of examination:

Name of the Test	Value	Grading
Examination ¹	2.5 credits	5/4/3/U
Predictive modeling assignment	2.5 credits	U/G
Descriptive modeling assignment	2.5 credits	U/G

¹ Determines the final grade of the course, which is issued only when all course units have been passed.

Course literature

The literature list for the course will be provided one month before the course starts.

Principal text:

Authors: Linoff G. and Berry M.

Title: Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management (third edition)

Publisher: Wiley (2011)

ISBN: 9780470650936 (paperback)

Additional texts:

Author: Demšar, J.

Title: Statistical comparisons of classifiers over multiple data sets.

Journal: Journal of Machine learning research, 7(Jan), 1-0. (2006)

3-5 additional research articles