COURSE SYLLABUS

Data mining med tillämpningar
Data Mining with Applications
4 credits (4 högskolepoäng)

Course code: DV2622
Main field of study: Computer Science
Disciplinary domain: Technology
Education level: Advanced level
Specialization: A1N - Second cycle, has only first-cycle sourse/s as entry requirements

Language of instruction: The language of instruction is English.
Applies from: 2022-08-29
Approved: 2022-03-01

1. Decision

This course is established by Dean 2021-12-03. The course syllabus is approved by Head of Department of Computer Science 2022-03-01 and applies from 2022-08-29.

2. Entry requirements

Admission to the course requires a Bachelor's degree in Computer Science. In addition, 7.5 credits in Programming and 5 credits in Machine Learning and English language equivalent to English 6.

3. Objective and content

3.1 Objective

The course aims to provide an introduction to theory and methods from data mining and knowledge discovery. It is intended for people with a background in computer science who have studied machine learning (ML) and artificial intelligence (AI) and want to expand their knowledge by learning more about the data mining pipeline, which includes data collection, cleaning, processing, analysis, and knowledge discovery.

3.2 Content

The course comprises the following topics, with intention to have at most a lecture per topic:
• Introduction to Data Mining: overview of the data mining process, its major building blocks, and challenges and applications
• Data Preparation: overview of the data cleaning, reduction and transformation, and dimensionality reduction
• Association Pattern Mining: introducing the problem of association pattern mining and identifying relationships between different attributes.
• Clustering Analysis: introducing variety of the clustering algorithms and their application, and different cluster validation methods
• Outlier Analysis: overview of outlier analysis and its application in different application domains, and outlier validation methods
• Mining Data Stream: overview of algorithms for stream mining and challenges related to steams such as high volume and concept drift

4. Learning outcomes

The following learning outcomes are examined in the course:

4.1 Knowledge and understanding

On completion of the course, the student will be able to:
• independently define and describe solvable and tractable data mining problems
• independently explain and summarize results from the application and evaluation of the studied problems

4.2 Competence and skills

On completion of the course, the student will be able to:
• independently identify the key components of the data mining pipeline and describe how they are related
• independently design and execute experiments to evaluate and compare data mining methods

4.3 Judgement and approach

On completion of the course, the student will be able to:
• independently evaluate and compare the performance of different data mining solutions using proper evaluation criteria
• independently analyze and interpret the experimental results from the evaluation of data mining solutions

5. Learning activities

Lectures are delivered via a video conference tool. In addition to the lectures' slides, a few complementary materials as a Jupyter notebook, will be provided, allowing the participants to dig into the concepts presented in the lecture and put them to practice. The students will demonstrate their knowledge in completing one assignment and a final project.

6. Assessment and grading

Modes of examinations of the course

Code	Module	Credits	Grade
2210	Written assignment	1 credits	GU
2220	Project assignment	3 credits	GU

The course will be graded G Pass, UX Insufficient, supplementation required, U Fail.

The information before a course occasion states the assessment criteria and make explicit in which modes of examination that the learning outcomes are assessed.

An examiner can, after consulting the Disability Advisor at BTH, decide on a customized examination form for a student with a long-term disability to be provided with an examination equivalent to one given to a student who is not disabled.

7. Course evaluation

The course evaluation should be carried out in line with BTH:s course evaluation template and process.

8. Restrictions regarding degree

The course can form part of a degree but not together with another course the content of which completely or partly corresponds with the contents of this course.

9. Course literature and other materials of instruction

Data Mining: The Textbook
Author: Charu C. Aggarwal
Publisher: Springer International Publishing Switzerland
Published: 2015, Number of Pages: 746
ISBN: 978-3-319-14141-1