Data Mining - Clustering and Association Analysis

Programme course

6 credits

Data Mining - Clustering and Association Analysis

TDDD41

Valid from: 2017 Spring semester

Determined by
Board of Studies for Computer Science
and Media Technology

Date determined
2017-01-25

Main field of study
Information Technology, Computer Science and Engineering, Computer Science

Course level
Second cycle

Advancement level
A1X

Course offered for
- Computer Science and Engineering, M Sc in Engineering
- Computer Science, Master's programme
- Information Technology, M Sc in Engineering
- Computer Science and Software Engineering, M Sc in Engineering
- Industrial Engineering and Management, M Sc in Engineering
- Industrial Engineering and Management - International, M Sc in Engineering

Entry requirements
Note: Admission requirements for non-programme students usually also include admission requirements for the programme and threshold requirements for progression within the programme, or corresponding.

Prerequisites
The course requires thorough knowledge in programming, discrete mathematics, data structures and algorithms and databases.

Intended learning outcomes

The course lays the foundation for professional work and research in which large amounts of data are explored, modified, modelled and assessed to uncover previously unknown patterns and trends. The course focuses on clustering and association analysis.
Having completed the course, the student should be able to:

- understand and be able to use important terminology in data mining
- understand and use the theory behind clustering and association analysis
- use knowledge about techniques for clustering and association analysis
- demonstrate insightful assessment of the quality of given data sets and the information content on which clustering and association analysis can be based
- use and evaluate tools for clustering and association analysis

Course content

Association analysis: concepts and methods related to frequent item sets and association rules such as Apriori principle, FP-growth, evaluation of association rules,

Clustering: concepts and methods related to partitional clustering methods, hierarchical clustering methods, density-based clustering methods, cluster evaluation

Teaching and working methods

The teaching comprises lectures and computer laboratory work. Lectures are devoted to theory, concepts and techniques. The techniques are practised in the computer laboratory work.

Examination

LAB1	Laboratory work	2 credits	U, G
TEN1	Written examination	4 credits	U, 3, 4, 5

Grades

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

Department

Institutionen för datavetenskap

Director of Studies or equivalent
Patrick Lambrix

Examiner
Patrick Lambrix

Course website and other links
http://www.ida.liu.se/~TDDD41

Education components
Preliminary scheduled hours: 26 h
Recommended self-study hours: 134 h

Common rules

Regulations (apply to LiU in its entirety)

The university is a government agency whose operations are regulated by legislation and ordinances, which include the Higher Education Act and the Higher Education Ordinance. In addition to legislation and ordinances, operations are subject to several policy documents. The Linköping University rule book collects currently valid decisions of a regulatory nature taken by the university board, the vice-chancellor and faculty/department boards.

LiU's rule book for education at first-cycle and second-cycle levels is available at http://styrdokument.liu.se/Regelsamling/Innehall/Utbildning_pa_grund-_och_avancerad_niva.