2025

Data mining (INFOMDM)

General information

Course ID
INFOMDM

Course type
Course

Credits
7.5 EC

Period(s)

- Period 1 (01-09-2025 until 07-11-2025)

Category / Level
M (Master)

Instruction language
English

Offered by
Utrecht University - Faculty of Science

Is mandatory for

- Data Science (DASC)

Is part of elective component in

- Artificial Intelligence (AINM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Energy Science (ENSM)
- Human-Computer Interaction (HCIM)
- Innovation Sciences (NWIM)
- Innovation Sciences (NWIM)
- Minor Applied Data Science Profile GSNS (PROFILE-ADS)

- Minor Complex Systems Profile GSNS as of 2021 (PROFILE-CS-21)
- Science-based Entrepreneurship Profile GSNS (PROFILE-SBE)

Course enrolment

Waiting list
You may be placed on a waiting list for this course. Final course enrolment will be determined later.

OSIRIS Student
Visit OSIRIS Student > Enrol / Course in order to register for this course.

Enrolment periods

- Period 1
  Timeslots
  B (TUE-morning, THU-afternoon)

  Enrolment period
  2 June 2025 09:00 until 20 June 2025 23:59

  Late enrolment
  18 August 2025 09:00 until 14 September 2025 23:59

  Unenrolment
  until 14 September 2025 23:59

  Period of education
  1 september 2025 until 7 november 2025

Course description

Course goals

After this course the student

- knows how important data-mining algorithms work, and what their properties are
- knows how to interpret the models and patterns produced by these data-mining algorithms
- knows how to match a data-analysis problem with the appropriate data mining algorithm(s)
- understands general concepts of data analysis, such as overfitting, the curse of dimensionality, the bias-variance decomposition and the VC-dimension
- knows how to perform a (comparative) data-mining experiment in a sound manner
- has gained experience with data mining algorithms through one or more assignments.

Assessment
The course is graded through:

- a digital exam
- one or more assignments
- homework exercises

For details we refer to the information provided on the course website.
To qualify for a repair of the final result the mark needs to be a 4 or 5, or “AANV”.
Prerequisites
It is assumed that the participant has knowledge of:

- algorithms and data structures.
- probability theory and statistics.
- calculus.
- linear algebra.

This is a required course for students in the Data Science (DASC) master program.

Content

Topics covered may include (content can vary somewhat from year to year):

- PAC learning and the VC-dimension
- the bias-variance decomposition
- classification tree algorithms
- bagging, (gradient) boosting and random forests
- graphical models (including Bayesian networks)
- frequent pattern mining (sets, sequences, trees)
- logistic regression
- text classification (with the multinomial naive Bayes model)
- classification for network link prediction
- Support Vector Machines

For further details see the course's website https://ics-websites.science.uu.nl/docs/vakken/mdm/

Course form
Lectures, tutorial sessions.

Literature
Selected book chapters, articles, and lecture notes.

Prerequisites and entry requirements

Entrance requirements course enrolment
You must meet the following requirements

- Assigned study entrance permit for the master

Instructional modes

Instructional modes

- Lecture

- Seminar
  Attendance requirement
  Yes

Tests

Tests

- Final result
  Test weight
  100

  Minimum grade

Lecturers

Course contact

- dr. A.J. Feelders

Examinator

- dr. A.J. Feelders

Lecturer

- dr. A.J. Feelders

Do you want to know more?

dr. A.J. Feelders can tell you more about this course.
Call +31 30 2533176 or send an e-mail to A.J.Feelders@uu.nl