Data Mining Lab Course SoSe 2026

Organization and Contents

L. Richter

Department of Computer Science
Technische Universität München

Premeeting, Wed, Feb 4th Room MI 01.09.034 4pm

Outline

Contents

Organization

Mission statement: At the end of the course you have gone through the whole path of data mining (knowledge discovery pipeline), from dataset preparation, detection of patterns and models, identifying and engineering features up to meaningful predictions. We provide you the glue to the theory.

Steps

This includes o.a. the following steps:

- Dataset search & Description
- Understanding the data & naïve Introspection
- Feature Construction & Selection
- Prediction & Evaluations

Tools

- Use scripting languages/scripts for data processing
- Optional: Database
- Use Weka/RapidMiner/Python/R for various processing steps:
  - exploratory data inspection
  - discretization
  - feature selection
  - machine learning algorithm implementations
- Own implementations of learning schemes are not intended
- Besides this you are free to use alternative tools of your choice

What we provide:

We provide infrastructure and support:

- A lab run in reverse class room style
- A wiki for the documentation of your work and communication with your peers and supervisors
- Mini-talks to selected topics from the theory
- Feedback to your questions and discussion of your findings

What we expect:

- Curiosity
- Self motivation
- Creativity to come up with own solutions
- Question yourself if you prefer clear instructions and directions → then **this** is not your course
- Readiness to retrieve (some) necessary background knowledge autonomously
- **Soft skills**: The ability to form and to collaborate in a team. No lonesome wolves!

Organization

- All of you are asked to suggest interesting datasets
- We select a dataset for each group
- There is one team of 3-4 people for one dataset
- You populate the wiki every week / weekly progress report
- You get substantial feedback every week in the plenary meetings
- Grading based on the group's success

Teaching Mode

- Whenever possible the weekly meetings are held in person
- There will be no hybrid solution, i.e. no additional video feed to the in person meetings
- I.e. students are supposed to come to Garching

Application

- Master students only
- Applicants (people present in this meeting or contacting me via mail) get a higher priority in the matching process
- No written CV, no motivation letter
- You ONLY express your interest by putting your name in the list today or sending an email to richter(at)in.tum.de if you could not make it to the pre-meeting
- There is no further selection than assigning a higher priority in the matching process.
- I have no resources to evaluate CVs, motivation letters or transcripts – Sorry!

What is left now?

- today we collect your name, matriculation and study program to identify you in the matching process
- please post the respective information here in the video conference chat
- after the matching is done we maintain a wait list
- some time ahead of the new semester we will send out an email and ask you to confirm your participation
- the timely response is **MANDATORY**
- not confirmed spots will be recycled for the wait list

Questions?