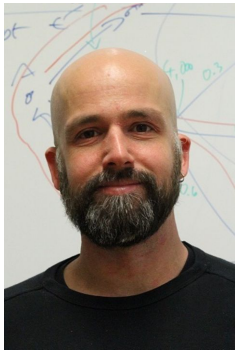
UNIVERSITÄT MANNHEIM

Data Mining II
Organization

Heiko Paulheim, Nicolas Heist

Hello



- Heiko Paulheim
- Professor for Data Science
- Research Interests:
  - Knowledge Graphs on the Web and their Applications
  - Data Quality and Data Cleaning on Knowledge Graphs
  - Using Knowledge Graphs in Data Mining
  - Societal Impact of Artificial Intelligence
- Consultation: Tuesdays, 9-10am
  - Please make an appointment via e-mail to Ms. Lermer
- Heiko will teach the lectures

Hello

- M.Sc. Andreea Iana
- Graduate Research Associate
- Research Interests:
  - Recommender systems
  - Information retrieval
  - Natural language processing
- eMail: andreea.iana@uni-mannheim.de
- Andreea will teach the exercises and co-supervise the challenge

Course Organization

- Lecture
  - addresses advanced data mining topics
  - builds on Data Mining I lecture contents!
- Project Work
  - we will take part in a challenge (to be determined)
  - you will work in teams
    - the two best performing teams submit their solutions
  - regular presentations of your approaches
  - paper and final presentation
- Exercise
  - weekly, covering lecture contents and previous challenge tasks

Requirements

- Final exam
  - 100 % written exam
  - project is not graded, but mandatory!
- Project work
  - work on challenge tasks
- Presentations
  - a few intermediate presentations
    - open questions, problems, current results (numbers!)
  - everybody has to present once during those presentations
- Final report
  - 10 pages
  - solutions, results, lessons learned

Schedule (tbc)

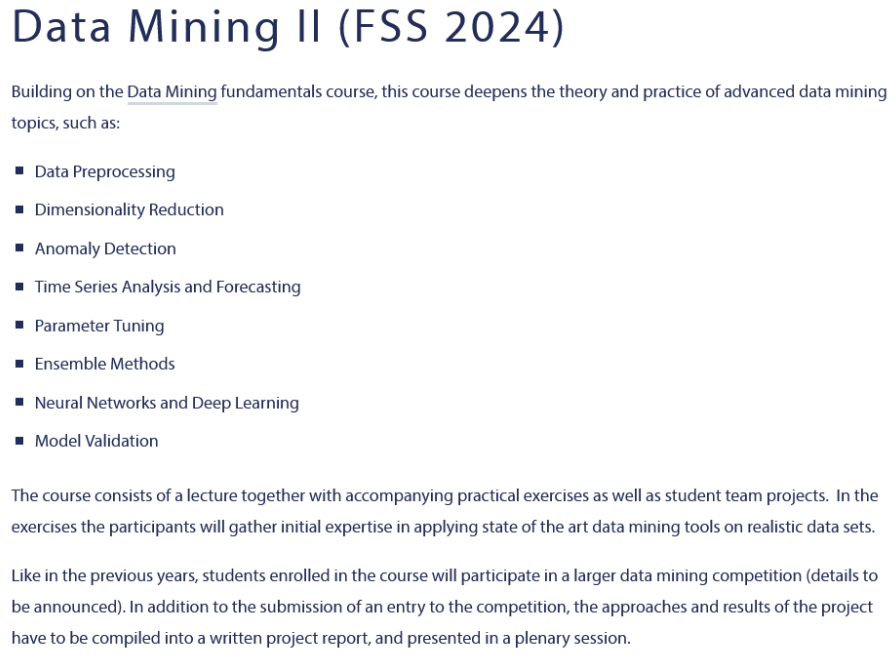
Week	**Exercise** (Monday)	**Lecture** (Tuesday)
12.2.	--	Introduction & Data Preprocessing
19.2.	Introduction & Data Preprocessing	--
26.2.	--	Ensembles
4.3.	Ensembles	Time Series
11.3.	Time Series	--
18.3.	--	Neural Networks & Deep Learning
25.3.	--	*Easter Break*
1.4.	--	*Easter Break*
8.4.	Neural Networks & Deep Learning	Anomaly Detection
15.4.	Anomaly Detection	Hyperparameter Tuning
22.4.	Hyperparameter Tuning	Model Verification
29.4.	Model Verification	Challenge Session
6.5.	--	Challenge Session
13.5.	--	Challenge Session
20.5.	--	Challenge Session

Deadlines

- Challenge deadlines will be announced
- Final report and first solutions due:
  - June 2nd, 2023
- Challenge solutions can still be updated

Course Organization

- Lecture Webpage: Slides, Announcements
  - https://www.uni-mannheim.de/dws/teaching/course-details/courses-for-master-candidates/ie-672-data-mining-2/
  - hint: look at version tags!
- Additional Material
  - ILIAS eLearning System, https://ilias.uni-mannheim.de/



Data Mining II (FSS 2024)

Building on the Data Mining fundamentals course, this course deepens the theory and practice of advanced data mining topics, such as:

- Data Preprocessing
- Dimensionality Reduction
- Anomaly Detection
- Time Series Analysis and Forecasting
- Parameter Tuning
- Ensemble Methods
- Neural Networks and Deep Learning
- Model Validation

The course consists of a lecture together with accompanying practical exercises as well as student team projects. In the exercises the participants will gather initial expertise in applying state of the art data mining tools on realistic data sets.

Like in the previous years, students enrolled in the course will participate in a larger data mining competition (details to be announced). In addition to the submission of an entry to the competition, the approaches and results of the project have to be compiled into a written project report, and presented in a plenary session.

Video Recordings of 2021

- We will provide lecture videos in ILIAS
- In case you prefer not to or cannot attend the lecture, you are advised to watch those lectures

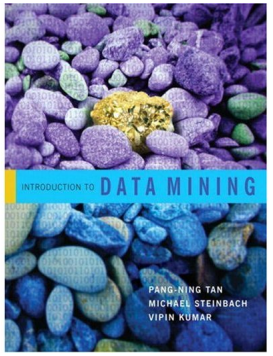
Exercises

- We offer two exercise groups
  - You only need to attend one
  - Room: A5, 6, C 013

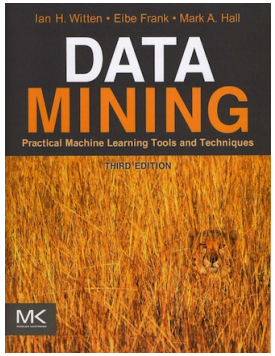
- Monday, 12:00 – 13:30
- Monday, 13:45 – 15:15

Literature & Slide Sources

- Pang-Ning Tan, Michael Steinbach, Vipin Kumar:
  Introduction to Data Mining,
  Pearson / Addison Wesley.
  - 10 copies in university library



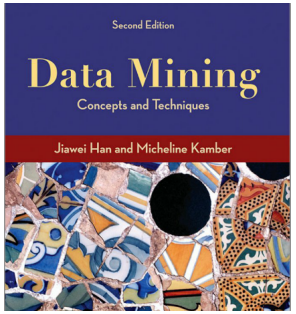
- Ian H. Witten, Eibe Frank, Mark A. Hall:
  Data Mining: Practical Machine Learning
  Tools and Techniques, 3rd Edition, Morgan Kaufmann.
  - several copies in university library



Literature & Slide Sources

- Gregory Piatetsky-Shapiro, Gary Parker:
  KDNuggets Data Mining course:
  http://www.kdnuggets.com/data_mining_course/

- Jiawei Han and Micheline Kamber:
  Data Mining – Concepts and Techniques
  - free e-book access via university library



Questions?