Data Mining I - Introduction to Data Mining

Summer

Introduction to Data Mining

(in German: Data Mining I - Introduction to Data Mining)

Module-ID: FIN-INF-120454

Abbreviation	Credit Points	Semester	Term	Duration	Language	Level
DM1_ENG	6	Summer	starting at 1.	1 Semester	english	Master

Link:	LSF
Limited accessability:	
Responsibility:	Myra Spiliopoulou
Lecturer:	Myra Spiliopoulou
Classes:	• Vorlesung DM1_ENG • Übung DM1_ENG **COURTESY TRANSLATION:** • Lecture class DM1_ENG • Exercise class DM1_ENG
Applicability in curriculum:	- M.Sc. DKE: Fundamentals of Data Science - M.Sc. DE: Methoden der Informatik - M.Sc. VC: Computer Science

Intended learning outcomes:
When successfully completing this course, the students

- comprehend the purpose of model induction from data
- are familiar with the behaviour of simple learning algorithms on tabular static data and can apply them to derive models
- comprehend the purpose, challenges and basic instruments for model evaluation
- can evaluate models
- comprehend the purpose of data engineering in preparation of data mining
- are familiar with basic tasks of data engineering

and have thus acquired skills they need in order to design themselves learning algorithms, and to build workflows for model evaluation.

Content:

- Underpinnings: basic forms of learning
- BLOCK 1 - supervised: Classification algorithms and evaluation procedured
- BLOCK 2 - unsupervised: Clustering algorithms and evaluation procedures
- BLOCK Evaluation: more on model evaluation and model comparison
- BLOCK Data Engineering: basic steps on engineering the data before they are input to the model learner

Workload:

- 28 hours - in presence : lecture class (2 hours per week)
- 28 hourse in presence: exercise class (2 hours per week)
- 124 hours: independent work (for lectures, exerices and for the exam preparation)

Pre-examination requirements:	Type of examination:	Teaching method / lecture hours per week (SWS):
A minimum number of points must be achieved during the course.	• Schriftliche Prüfung - Klausur (120 min) **COURTESY TRANSLATION:** • Written exam of the form 'Klausur' (120 min)	• Vorlesung (2 SWS) • Übung (2 SWS) **COURTESY TRANSLATION:** • Lecture (2 hours per week of the semester) • Exercise (2 hours per week of the semester)

Prerequisites according to examination regulations:	Recommended prerequisites:
keine	• Databases • Data management • Basics in statistics

Media:	Literature:
	Book M: The book of the course is *Introduction to Data Mining* by Pan-Ning Tan, Michael Steinbach, Anuj Karpatne and Vipin Kumar. PEARSON, 2019 (2nd edition) • Chapters 2 and 3 • Parts of Chapters 4, 6, 7 • Parts of Chapters 8 and 10 Overview of the book under https://www-users.cs.umn.edu/~kumar001/dmbook/index.php including slide collections etc To read the full textbook (electronic version): Find the ebook from the library here, follow the direct link, and log in -> Shibboleth login -> Otto-von-Guericke University -> Log in -> read. *Note:* Some slides come from the 1st edition. They concern content that did not change between the two editions. **Book D:** *Data Preprocessing in Data Mining* by Salvador Garcia, Julian Luengo and Fransisco Herrera, SPRINGER International Publishing Switzerland, 2015**** This book is used in the Block 'Data Engineering' (Block 'Data' for short). This book is available as pdfs from our library. The chapters of relevance are mentioned in the description of the course block.****

Comments:

This module is intended for students in low semesters of 4-semester-long master degrees and assumes a limited background in data science. It is not appropriate for students with a bachelor degree in data science.

Revision #1
Created 2026-10-20 17:13:24 UTC by Fakultät für Informatik
Updated 2026-10-20 17:13:24 UTC by Fakultät für Informatik