# Data Mining – Einführung in Data Mining

**Winter**

## Einführung zu Data Mining für Bachelorstudiengänge

(engl. Data Mining for Bachelor Students - An Introduction (German course) )

### Modulnummer: FIN-INF-110302

| Kürzel | CP | Semester | Fachsem. | Dauer | Sprache | Niveau |
|---|---|---|---|---|---|---|
| DM4BA | 5 | Winter | 3. | 1 Semester | deutsch | Bachelor |

| | |
|---|---|
| **Link zum LSF:** | LSF |
| **Zugang beschränkt:** | false |
| **Verantwortung:** | Myra Spiliopoulou |
| **Dozent:in:** | Myra Spiliopoulou |
| **Lehrveranstaltungen:** | • Vorlesung *DM4BA*<br>• Übung *DM4BA* |
| **Verwendbarkeit:** | - B.Sc. INF: Informatik - Wahlpflicht<br>- B.Sc. INF: Studienprofil: Künstliche Intelligenz<br>- B.Sc. CV: Informatik - Wahlpflicht<br>- B.Sc. INGINF: Informatik - Wahlpflicht<br>- B.Sc. WIF: Verstehen und Gestalten - Wahlpflicht<br>- B.Sc. WIF: Gestalten und Anwenden - Wahlpflicht<br>- B.Sc. INF (bilingual): Informatik - Wahlpflicht |

**Angestrebte Lernergebnisse:**
Die Studierenden:

- verstehen, wozu Data Mining gut ist
- verstehen, wie eine Klassifikationsaufgabe aussieht und wie eine Clustering-Aufgabe aussieht, und können zwischen den zwei Typen von Aufgaben unterscheiden

- verstehen, wie einfache Klassifikationsmethoden funktionieren und können diese für Probleme anwenden
- verstehen, wie einfache Clustering-Methoden funktionieren und können diese für Probleme anwenden
- verstehen die Herausforderungen der Evaluation von Modellen
- können Evaluationsvorgänge gestalten und durchführen
- verstehen, warum Data Engineering vor dem Aufruf von Lernverfahren nötig ist
- können einfache Instrumente der Datenaufbereitung anwenden und ihre Ergebnisse evaluieren

und haben damit notwendige Grundlagen erworben, um einen einfachen Vorgang zum Ableiten eines Modells und zu dessen Evaluation zu gestalten.

**Inhalt:**

- Klassifikation: Lernverfahren und Evaluationsvorgänge
- Clustering: Lernverfahren und Evaluationsvorgänge
- mehr zu Methoden für die Evaluation von Modellen
- Data Engineering: Datenaufbereitungsaufgaben, -methoden und -evaluationsvorgänge

**Arbeitsaufwand:**

- 28 Stunden Vorlesung (Präsenz)
- 28 Stunden Übung (Präsenz)
- 94 Stunden Eigensändiges Arbeit zur Vor- und Nachbereitung der Vorlesung und der Übung, sowie zur Prüfungsvorbereitung

| Prüfungsvorleistungen: | Studien-/Prüfungsleistungen: | Lehrform / SWS: |
|---|---|---|
| Votierung - min Anzahl von Punkten durch Aufgaben in den Übungen | Schriftliche Prüfung - Klausur (120 min) | • Vorlesung (2 SWS)<br>• Übung (2 SWS) |

| Voraussetzungen nach Prüfungsordnung: | Empfohlene Voraussetzungen: |
|---|---|
| keine | Datenbanken |

| Medienformen: | Literatur: |
|---|---|

- Zu Data Mining Methoden und zu Evaluationsansätzen:
  - Pan-Ning Tan, Michael Steinbach, Anuj Karpatne, Vipin Kumar (2019). *Introduction to Data Mining*, 2nd edition. Pearson – Ausschnitte aus Kpt 3, 5, 7
  - Ian Witten & Eibe Frank (2005). *Data Mining – Practical machine learning tools and techniques*. 2nd edition.

    Elsevier – Ausschnitte aus Kpt 4 zu Naive Bayes
  - Zu Data Engineering:
    - Salvador Garcia, Julian Luengo, Francisco Herrera (2015). *Data Preprocessing for Data Mining*. Springer – Ausschnitte aus Kpt 3, 4, 7, evtl auch aus Kpt 2
    - und eventuell für Zeitreihen:
      - Parmezan, A. R. S., Souza, V. M. A., and Batista, G. E. A. P. A. (2019). *Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model*. Information Sciences, DOI: 10.1016/j.ins.2019.01.076
      - Cismondi, F. C., Fialho, A., Vieira, S., Reti, S., Sousa, J., and Finkelstein, S. (2013). *Missing data in medical databases: Impute, delete or classify?* Artificial Intelligence in Medicine, DOI: 10.1016/j.artmed.2013.01.003
  - Weitere Literatur wird in den einzelnen Sessions aufgelistet.

## Hinweise:

Revision #1
Created 2026-10-20 16:54:05 UTC by Fakultät für Informatik
Updated 2026-10-20 16:54:05 UTC by Fakultät für Informatik