



Johann Wolfgang Goethe-Universität Frankfurt am Main

Data Mining

[University Home](#)

[University
Overview](#)

[Suche](#)

[Dept. of
Computer Science](#)

[Databases and
Information
Systems](#)

**Vorlesung - Sommer 2002: Freitags, 10-12 Uhr, Magnus-Hörsaal
Dr. Christoph Schommer**

A. Veranstaltungstermine und Organisatorisches

- 12. April 2002: [Einführung](#) in Data Mining
- 19. April 2002: [Daten, Datenhaltung und Datendiskussion](#)
- 26. April 2002: [Data Preprocessing](#)
- 03. Mai 2002: [Statistische Grundlagen](#)
- 10. Mai 2002: [Generierung assoziativer Regeln](#) - Teil 1
- 17. Mai 2002: [Generierung assoziativer Regeln](#) - Teil 2
- 24. Mai 2002: [Suche nach sequentiellen Mustern](#) bzw. [Suche nach Ähnlichkeiten zwischen Zeitreihen](#)
- 28. Mai 2002: [Profilierung mit Hilfe von Clustering](#)
- 07. Juni 2002: [Predictive Modelling am Beispiel von Decision Trees](#)
- 14. Juni 2002: [Visualisierung und Interpretation](#)
- 21. Juni 2002: [Anwendungen](#) aus der Praxis
- 28. Juni 2002: Klausurvorbereitung: Wichtigste Aspekte:
 - [Übungsfrage 1](#)
 - [Übungsfrage 2](#)
 - [Übungsfrage 3](#)
 - [Übungsfrage 4](#)
 - [Übungsfrage 5](#)
 - [Übungsfrage 6](#)
 - [Übungsfrage 7](#)
 - [Übungsfrage 8](#)
- **05. Juli 2002 (Magnus Hörsaal, 10-12): [Klausur](#)**
Anmeldung bitte entweder in der Vorlesung oder über Email an [cs](#)
(Angabe von Namen, Matrikelnummer, Semesterzahl, Geburtsdatum)

B. Begleitmaterialien

Praktische Data Mining-Aufgabe (Quelle: [Data Mining Cup 2002](#))

- [Aufgabenstellung und Szenario](#)
- [Kurzerklärung der Merkmale](#)
- [Ausführliche Beschreibung der mikrogeografischen Variablen](#)
- [Trainings-, Testdaten](#)
- [Anwendungsdaten](#)

Zusätzliches Material

- [Database Skript](#) zum Kreieren der Datenbanktabelle und -view; Einlesen der Daten nach DB2

- Skript zum Pivotisieren der Datenmenge: [pivot.awk](#)
- Datenmenge [ecustomer](#)
- Pivotisierte Datenmenge [pivot](#)
- [Large Itemsets](#)
- [Assoziationsregeln](#) mit min Support=1% und min Confidence=4% und Rule_length=2
- [Allgemeine Infos - Assoziationsregeln](#)
- [Multilayer Perceptron, von C. Borgelt, Uni Magdeburg](#)
- Roche-Lexikon für Medizin: [Adrenatitis](#)

C. Werkzeuge



Datenbank: DB2 Universal Database



Data Mining: Intelligent Miner for Data

D. Referenzen

- I. Witten, E. Frank: Data Mining. Morgan Kaufman Publishers.
- M. Lusti: Data Mining and Data Warehousing. Springer Verlag.
- IBM Redbook Series. Free download at [www.redbooks.ibm.com](#)
- KDnuggets --> [www.kdnuggets.com](#)

E. Universitäre und industrielle Kontakte

- Biologie (Inst. für Mikrobiologie): [Prof. Entian](#)
- Datenbanken (Inst. für Informatik): [Prof. Zicari](#)
- Medizinische Informatik (Fachbereich Medizin): [Prof. Giere](#)
- IBM Entwicklung GmbH, Böblingen: [Gerd Piel](#)
- IBM Informationssysteme GmbH, Stuttgart: [Peter Starting](#)
- pepper technologies GmbH, Berlin: [Dr. Peter Gentsch](#)

F. Zertifizierungen (*IBM certified ...: in Planung*)

- DB2 Universal Database V7.2, V8
- DB2 Intelligent Miner for Data V6.1.1, V8

G. Sonstiges

- Sprechstunde: Freitags, n.V. (Zi. 107, Zeppelinallee 6-8)
- [Folgeveranstaltung@Winter 2002/2003](#): Anwendung von Data Mining (Praktikum)

Last Update: April 20, 2002