

# Lehrveranstaltung

## DDML - Data Mining

---

Version: 1 | Letzte Änderung: 12.02.2021 13:46 | Entwurf: 0 | Status: vom verantwortlichen Dozent freigegeben

### ^ Allgemeine Informationen

<b>Langname</b>	Data Mining
<b>Anerkennende LModule</b>	<a href="#">DML_BaET</a>
<b>Verantwortlich</b>	Prof. Dr. Beate Rhein Professor Fakultät IME
<b>Niveau</b>	Bachelor
<b>Semester im Jahr</b>	Wintersemester
<b>Dauer</b>	Semester
<b>Stunden im Selbststudium</b>	78
<b>ECTS</b>	5
<b>Dozenten</b>	Prof. Dr. Beate Rhein Professor Fakultät IME
<b>Voraussetzungen</b>	Aus Mathematik 1 und 2 wird die Fähigkeit benötigt, mathematische Modelle aufzustellen, sowie Kenntnisse der Differentialrechnung und der Linearen Algebra.
<b>Unterrichtssprache</b>	deutsch
<b>separate Abschlussprüfung</b>	Ja

## Abschlussprüfung

### Details

Abhängig von der Anzahl der Teilnehmer:

Bei wenigen Teilnehmern Kombination aus Klausur oder mündlicher Prüfung und Bewertung des Mini-Projektes

Bei vielen Teilnehmern Prüfung durch Klausur; Mini-Projekt als Voraussetzung zur Teilnahme an der Klausur

In der Klausur bzw. mündlichen Prüfung werden die Methoden, Vorgehensweisen, Fallstricke und gesetzliche Grundlagen des Data Mining geprüft.

Im Mini-Projekt wird die Fähigkeit zu selbständigem eigenverantwortlichem Handeln und der Umgang mit geeigneter Software abgeprüft.

## Mindeststandard

Grundkenntnisse der generellen Vorgehensweise beim Data Mining, der behandelten Verfahren und ihre Grenzen

## Prüfungstyp

Abhängig von der Anzahl der Teilnehmer:

Bei wenigen Teilnehmern Kombination aus Klausur oder mündlicher Prüfung und Bewertung des Mini-Projektes

Bei vielen Teilnehmern Prüfung durch Klausur; Mini-Projekt als Voraussetzung zur Teilnahme an der Klausur

In der Klausur bzw. mündlichen Prüfung werden die Methoden, Vorgehensweisen, Fallstricke und gesetzliche Grundlagen des Data Mining geprüft.

Im Mini-Projekt wird die Fähigkeit zu selbständigem eigenverantwortlichem Handeln und der Umgang mit geeigneter Software abgeprüft.

# ^ Vorlesung / Übungen

## Lernziele

---

### Kenntnisse

Einführung in eine geeignete Software, z.B. Python

Einführung in deskriptive Statistik und evtl. auch Wahrscheinlichkeitsrechnung

Überwachtes Lernen:

- Klassifikationsverfahren: Ablauf, Performanzmaße, Anwendung eines Verfahrens des instanzbasierten Lernen, z.B. k-nearest-neighbor und eines

Verfahrens des modellbasierten Lernen, z.B. Entscheidungsbäume

- evtl. Regressionsanalyse: über maschinelles Lernen und klassisch

Unüberwachtes Lernen:

- Clusteranalyse: k-means, evtl. auch DBSCAN

Preprocessing der Daten:

- Behandlung von beschädigten / fehlenden Daten

- Ausreißer oder Noise - Problematik

- Skalierung

- Visualisierung der Daten

- evtl. Dimensionsreduzierung

- Beurteilung der Qualität der Daten

- evtl. verschiedene Arten von Datensätzen betrachten, Bezug zu NoSql-Datenbanken herstellen, Ausblick auf aktuelle Forschung, z.B. Bilderkennung,

Natural Language Processing, Reinforcement Learning

---

### Fertigkeiten

zu Aufgabenstellungen geeignete Methode und Gesamtvorgehensweise nennen und anwenden können

geeignetes Performanzmaß wählen und beurteilen

Datenschutzrichtlinien anwenden

## Aufwand Präsenzlehre

<b>Typ</b>	<b>Präsenzzeit (h/Wo.)</b>
Vorlesung	2
Übungen (ganzer Kurs)	0
Übungen (geteilter Kurs)	2
Tutorium (freiwillig)	0

## Separate Prüfung

keine