

Modul

DM - Data Mining

Bachelor Technische Informatik 2020

Version: 1 | Letzte Änderung: 27.09.2019 12:39 | Entwurf: 0 | Status: vom Modulverantwortlichen freigegeben | Verantwortlich: Rhein

^ Allgemeine Informationen

Anerkannte Lehrveranstaltungen	DM_Rhein
Modul ist Bestandteil der Studienschwerpunkte	SMS - Smart Systems IOT - Internet of Things
Dauer	1 Semester
ECTS	5
Zeugnistext (de)	Data Mining
Zeugnistext (en)	Data Mining
Unterrichtssprache	deutsch
abschließende Modulprüfung	Ja

Modulprüfung

Benotet	Ja
Frequenz	Jedes Semester

Prüfungskonzept

Abhängig von der Anzahl der Teilnehmer:

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

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

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

^ Allgemeine Informationen

Inhaltliche Voraussetzungen

MA1 - Mathematik 1	mathematische Modelle verstehen und aufstellen Differentialrechnung
MA2 - Mathematik 2	Funktionen mit mehreren Veränderlichen anwenden Lineare Algebra: Matrizen aufstellen und mit ihnen rechnen
PI1 - Praktische Informatik 1	Grundlagen der Programmierung beherrschen

Kompetenzen

Kompetenz	Ausprägung
fachliche Probleme abstrahieren und formalisieren	Vermittelte Kompetenzen
Konzepte und Methoden der Informatik, Mathematik und Technik kennen und anwenden	Vermittelte Kompetenzen
Informationen beschaffen und auswerten; Technische Zusammenhänge darstellen und erläutern	Vermittelte Kompetenzen
Gesellschaftliche und ethische Grundwerte anwenden	Vermittelte Voraussetzungen für Kompetenzen
Befähigung zum lebenslangen Lernen	Vermittelte Voraussetzungen für Kompetenzen

^ Vorlesung / Übungen

Exemplarische inhaltliche Operationalisierung

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

Separate Prüfung

keine