

Lehrveranstaltungshandbuch DM

Data Mining

Version: 1 | Letzte Änderung: 27.09.2019 12:52 | Entwurf: 0 | Status: vom verantwortlichen Dozent freigegeben

– Allgemeine Informationen

Langname	Data Mining
Anerkennende LModule	<u>DM BaTIN</u>
Verantwortlich	Prof. Dr. Beate Rhein Professor Fakultät IME
Gültig ab	Sommersemester 2022
Niveau	Bachelor
Semester im Jahr	Wintersemester
Dauer	Semester
Stunden im Selbststudium	78
ECTS	5
Dozenten	Prof. Dr. Beate Rhein Professor Fakultät IME
Voraussetzungen	Aus Mathematik 1 und 2 wird die Fähigkeit benötigt, mathematische Modelle aufzustellen, sowie Kenntnisse der Differentialrechnung und der Linearen Algebra.
Unterrichtssprache	deutsch
separate Abschlussprüfung	Ja

Literatur

A. Geron: Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme, Heidelberg, o'Reilly Verlag 2017, 978-3960090618

S. Raschka, V. Mirjalili: Machine Learning mit Python und Scikit-Learn und TensorFlow: Das umfassende Praxis-Handbuch für Data Science, Predictive Analytics und Deep Learning, mitp Verlag, 2018, 978-3-958457331

J. Frochte, Jörg: Maschinelles Lernen, München, Carl Hanser Verlag GmbH & Co. KG, 2018, eBook ISBN: 978-3-446-45705-8, Print ISBN: 978-3-446-45291-6

A. Müller: Einführung in Machine Learning mit Python: Praxiswissen Data Science, Heidelberg, o'Reilly Verlag 2017, eBook: 978-3-96010-111-6

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

andere summarische Prüfungsform

– Vorlesung / Übungen

Lernziele

Zieltyp	Beschreibung
Kenntnisse	<p>Einführung in eine geeignete Software, z.B. Python</p> <p>Einführung in deskriptive Statistik und evtl. auch Wahrscheinlichkeitsrechnung</p> <p>Überwachtes Lernen:</p> <ul style="list-style-type: none">- 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 <p>Unüberwachtes Lernen:</p> <ul style="list-style-type: none">- Clusteranalyse: k-means, evtl. auch DBSCAN <p>Preprocessing der Daten:</p> <ul style="list-style-type: none">- 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 <p>Ausblick auf aktuelle Forschung, z.B. Bilderkennung, Natural Language Processing, Reinforcement Learning</p>
Fertigkeiten	<p>zu Aufgabenstellungen geeignete Methode und Gesamtvorgehensweise nennen und anwenden können</p> <p>geeignetes Performanzmaß wählen und beurteilen</p> <p>Datenschutzrichtlinien anwenden</p>

Besondere Voraussetzungen

keine

Begleitmaterial

Skript bzw. Foliensatz
Aufgabenstellungen
(voraussichtlich in Skript integriert)
Mini-Projekt-Aufgabe mit Datensatz

Separate Prüfung

Nein

Aufwand Präsenzlehre

Typ	Präsenzzeit (h/Wo.)
Vorlesung	2

Übungen (ganzer Kurs)	0
-----------------------	---

Übungen (geteilter Kurs)	2
--------------------------	---

Tutorium (freiwillig)	0
-----------------------	---