

21 Data Science	
Data Science	
Semester	Wahlpflichtbereich
Credit Points	5
Pflicht/ Wahlpflicht	Wahlpflicht
Häufigkeit des Angebotes/ Verwendbarkeit	Jedes Semester nach Bedarf der Hochschulen im VFH-Verbund
Modulverantwortliche(r)	Prof. Dr. Stefan Edlich, Beuth Hochschule für Technik Berlin
Lerngebiet	Informatik und Mathematik / Statistik
Teilnahmevoraussetzungen	Sichere Anwendung von Hochsprachen wie Java, sowie sichere Kenntnisse aus den Mathematikkursen des Bachelor-Studiengangs und der Wahrscheinlichkeitsrechnung.
Lernergebnisse	<p>Nach dem erfolgreichen Abschluss des Studienmoduls, sind die Studierenden in der Lage:</p> <ul style="list-style-type: none"> • die Grundlagen zu Vektorräumen, Matrizen und Wahrscheinlichkeit zu verstehen • Design und Struktur von Big Data Analytics Anwendungen zu entwerfen • Datensätze zu analysieren • Machine Learning (ML) -Modelle zu evaluieren und zu bewerten • Grundlagen wie CAP Theorem und NoSQL zu erklären • die Sprachen R und Python anzuwenden um bspw. Daten zu importieren, Data cleaning durchzuführen und Data Frames umzuwandeln • Bibliotheken zur Erstellung von Machine Learning Models anzuwenden • Fragestellungen des Machine Learning in Statistik und Informatik zu verbinden und zu implementieren • alle ML-Methoden darzustellen • die Arbeitsweise der wichtigsten ML-Algorithmen zu erklären
Prüfungsvorleistung	Einsendeaufgabe, Hausarbeit/Projekt/Übung
Medien-/ Lernform	Multimedial aufbereitetes Online-Studienmodul zum Selbststudium mit zeitlich parallel laufender Online-Betreuung (E-Mail, Foren, Chat, Webkonferenzen, Einsendeaufgaben u. a.)
Arbeitsaufwand	<p>Prüfung: 30 Minuten</p> <p>Selbststudium: 96 h</p> <p>Betreutes Lernen: 30 h</p> <p>Vorbereitung PVL: 24 h</p>

Präsenzinhalte	Klärung inhaltlicher Fragen
Prüfungsform	mündliche Prüfung/ Referat (30 min.)
Literatur	„Machine Learning“, Kevin P. Murphy ISBN-13: 978-0262018029 „Doing Data Science“ O’Neill & Schutt, ISBN-13: 978-1449358655
weitere Hinweise	Dieses Modul wird auf Englisch angeboten

Studieninhalte
00 Introduction
01 Linear Algebra
02 Python für Data Science
03 R and Julia
04 Machine Learning Introduction
05 Regression
06 Instance Based Methods
07 Decision Trees
08 Clustering
09 Data Preparation
10 Datasets
11 Hands-On
12 Data Visualization
13 Scalable Big Data Analytic Engines
14 Deep Learning and Neuronal Networks
15 Reinforcement Learning applied on Games