

Data Science & Big Data (W3M20018)

Data Science & Big Data

FORMALE ANGABEN ZUM MODUL

MODULNUMMER	VERORTUNG IM STUDIENVERLAUF	MODULDAUER (SEMESTER)	MODULVERANTWORTUNG	SPRACHE
W3M20018	-	1	Prof. Dr. Christoph Sturm	Deutsch/Englisch

EINGESETZTE LEHRFORMEN

LEHRFORMEN	LEHRMETHODEN
Vorlesung, Übung	Lehrvortrag, Diskussion

EINGESETZTE PRÜFUNGSFORMEN

PRÜFUNGSLEISTUNG	PRÜFUNGSUMFANG (IN MINUTEN)	BENOTUNG
Kombinierte Modulprüfung - Laborarbeit (Gewichtung 50%) und Klausur (Gewichtung 50%)	Siehe Prüfungsordnung	ja

WORKLOAD UND ECTS-LEISTUNGSPUNKTE

WORKLOAD INSGESAMT (IN H)	DAVON PRÄSENZZEIT (IN H)	DAVON SELBSTSTUDIUM (IN H)	ECTS-LEISTUNGSPUNKTE
150	48	102	5

QUALIFIKATIONSZIELE UND KOMPETENZEN

FACHKOMPETENZ

Die Studierenden erhalten einen fundierten Überblick über die Anwendungsgebiete von Big Data. Sie verstehen Big Data Architekturen und können sich mit deren Einsatzgebieten, Stärken und Schwächen reflektiert auseinandersetzen. Die Studierenden können Big Data Processing Frameworks zur Datenverarbeitung selbstständig auf betriebliche Problemstellungen übertragen und reflektieren. Sie verstehen die Funktionsweise spezifischer Algorithmen zur Analyse großer Datenmengen und können die Erfolgswahrscheinlichkeit ihres Einsatzes bewerten.

METHODENKOMPETENZ

Die Studierenden können die üblichen Methoden und Techniken von Big Data und Data Science in ihrem Unternehmen selbstständig einsetzen und reflektiert weiterentwickeln. Sie haben die Fähigkeit die Grenzen der Methoden für das Vorhaben ihres Unternehmens einzuschätzen. Die Studierende können die Methode der datengetriebenen Informations- und Wissensgewinnung zur Lösung von Problemen oder Beantwortung von Frageneinsetzen und diese auf beliebige Einsatzgebiete adaptieren. Dies beinhaltet die Auswahl der richtigen Algorithmen für die jeweilige praktische Fragestellung, die für den jeweiligen Algorithmus erforderliche Datenstrukturierung und die Bewertung des Ergebnisses.

PERSONALE UND SOZIALE KOMPETENZ

-

ÜBERGREIFENDE HANDLUNGSKOMPETENZ

Die Studierenden verstehen die datengetriebene und datenzentrierte Entscheidungsfindung.

LERNEINHEITEN UND INHALTE

LEHR- UND LERNEINHEITEN	PRÄSENZZEIT	SELBSTSTUDIUM
Data Science & Big Data	48	102

LERNEINHEITEN UND INHALTE

LEHR- UND LERNEINHEITEN

PRÄSENZZEIT

SELBSTSTUDIUM

Big Data
- Small vs. Big Data
Chancen und Auswirkungen von Big Data
- Big Data Use Cases
- Communication Cost Model
- Architekturen (Lambda-/Kappa-Architekturen)
- Datenverarbeitung großer Datenmengen
Cloud (IaaS/PaaS/SaaS)
Big Data Transfer / Storage
- Big Data Processing Frameworks
Code Pushdown
Batch Processing
Stream Processing

Data Science
- Definition / Einführung Data Science
CRISP-DM Prozess / Knowledge Discovery (KDD) Prozess
Data Mining Basistechniken
Data Science Use Cases
- Datenvorverarbeitung
Datenarten, Datenverständnis, Datenvisualisierung
Datenqualität
Datenintegration
Datenreduktion
Datentransformation
- Analyse großer Datenmengen
Supervised Segmentation
Overfitting
Linear / Logistic Regression
Support Vector Machines (SVM)
Similarity Search and Distance Measures
Clustering (KMeans)
Dimensionality Reduction (SVD, PCA)
Large Scale Machine Learning
Evaluation der Modelle

BESONDERHEITEN

Die Kombinierte Prüfung besteht aus den Prüfungsteilen Laborbericht und Klausur (Gewichtung: 50 %).

VORAUSSETZUNGEN

DB Kenntnisse, Programmierkenntnisse

LITERATUR

Cleveland, William S. (1993): Visualizing Data. Hobart Press.
Cleveland, William S. (1994): The Elements of Graphing Data. Hobart Press.
Few, Stephen (2006): Now You See It: Simple Visualization Techniques for Quantitative Analysis. Analytics Press.
Leser, Ulf und Naumann, Felix (2006): Informationsintegration. dpunkt.verlag.
Leskovec, Jure, Rajaraman, Anand und Ullman, Jeff (2014): Mining of Massive Datasets. 2. Auflage. Cambridge University Press.
Provost, Foster und Fawcett, Tom (2013): Data Science for Business. O'Reilly.
Schmarzo, Bill (2013): Big Data: Understanding How Data Powers Big Business. Wiley.
Segaran, Toby und Hammerbacher, Jeff (2009): Beautiful Data. O'Reilly.
Tufte, Edward R (2001): Visual Display of Quantitative Information. 2. Auflage. Bertrams.