

Modulcode (1.)	Modulbezeichnung (2.)	Zuordnung (3.)
MAI2550	Advanced Analytics (AA)	
	Studiengang (4.)	Master Angewandte Informatik
	Fakultät (5.)	Gebäudetechnik und Informatik

Modulverantwortlich (6.)	Prof. Dr. Marcel Spehr
Modulart (7.)	Wahl
Angebotshäufigkeit (8.)	WS
Regelbelegung / Empf. Semester (9.)	MA2
Credits (ECTS) (10.)	5 CP
Leistungsnachweis (11.)	PrP(N) SL
Unterrichtssprache (12.)	Deutsch
Voraussetzungen für dieses Modul (13.)	BAI1010: Mathematik 1 BAI1040: Grundkonzepte der Programmierung BAI0105: Datenbanken 1 BAI2010: Mathematik 2 BAI2050: Datenbanken 2 BAI3050: Statistik/Stochastik BAI6320: Data Analytics
Modul ist Voraussetzung für (14.)	
Moduldauer (15.)	1 Semester
Notwendige Anmeldung (16.)	-
Verwendbarkeit des Moduls (17.)	-

Lehrveranstaltung (18.)	Dozent/in (19.)	Art (20.)	Teilnehmer (maximal) (21.)	Anzahl Gruppen (22.)	SWS (23.)	Workload	
						Präsenz (24.)	Selbststudium (25.)
1 Advanced Analytics	Spehr	S	15	1	4	60	65
Summe					4	60	65
Workload für das Modul (26.)						125	

Qualifikationsziele	Die Studierenden können... • moderne Analytics Methoden benennen und verwenden • In modernen verteilten Systemen (Cluster, Cloud) rechenintensive Aufgaben geeignet formulieren und realisieren • Fokus liegt auf anwendungsnahen, integrierten Systemen mit zentraler KI-Komponente • selbstständig ausgehend von einer analytischen Problemstellung geeignete Werkzeuge und Methoden auswählen, sich diese erschließen und zielgerichtet einsetzen • selbstständig aktuelle Entwicklungen des Themas verfolgen und im Projektalltag moderne Methoden einsetzen • das Zusammenspiel von Vorhersagedomäne, Datencharakteristik, Bewertungsmetrik und Geschäftsentscheidungen einschätzen • den Hintergrund und die Wertschöpfungskette analytischer Daten beschreiben • analytische Problemstellungen einordnen und Methoden sowie Werkzeuge lösungsorientiert einsetzen • die Tools des Python Data Science Stacks für die Problemlösung korrekt verwenden • den gesellschaftlichen Einfluss von Datensammlungs- und Auswertesystemen erkennen und eigene Tätigkeiten kritisch, gesamtgesellschaftlich reflektieren
Inhalte	• Methodisch kommen moderne Algorithmen der KI zum Einsatz für die Lösung von Clustering-, Regressions- und Klassifikationsproblemen • Umsetzungswerkzeug für die Implementierung ist die Programmiersprache Python mit dem dedizierten Data Science Stack • Verwendung von Python Bibliotheken für Datenanalyse und Umsetzung der KI-Architekturen in verteilten Systemen (bspw. Spark) • Besonderheiten von Streaming Systemen unter Aspekten des KI-Trainings • Aktuelle Methoden und Werkzeuge für die Analysekette von deskriptiver, diagnostischer, prädiktiver und präskriptiver Verarbeitung • Vorgehen um ausgehend von einer speziellen Datencharakteristik und einem konkreten Vorhersageziel Bewertungsmetriken und Algorithmik für die Problemlösung zu definieren • Methoden für Datenanonymisierung und diskriminierungsfreies Modelltraining
Vorleistungen und Modulprüfung	• Vorleistungen: ◦ Keine • Modulprüfung ◦ Seminarfacharbeit zu einem vorgegebenen Analytics Problemfall mit Präsentationen
Literatur	• Jake VanderPlas, Python Data Science Handbook: Essential Tools for working with Data • Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython

	• Joel Grus, Data Science from Scratch: First Principles with Python • Peter Bruce et. Al., Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python • Alex Reinhard, Statistics Done Wrong: The Woefully Complete Guide, 2015 • Sandy Ryza et al, Advanced Analytics with Spark, 2017 • Mahmoud Parsian, Data Algorithms: Recipes for Scaling Up with Hadoop and Spark, 2015
--	--