

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

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

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

<b>Qualifikationsziele</b>	<p>Die Studierenden können...</p> <ul style="list-style-type: none"> <li>• moderne Analytics Methoden benennen und verwenden</li> <li>• In modernen verteilten Systemen (Cluster, Cloud) rechenintensive Aufgaben geeignet formulieren und realisieren</li> <li>• Fokus liegt auf anwendungsnahen, integrierten Systemen mit zentraler KI-Komponente</li> <li>• selbstständig ausgehend von einer analytischen Problemstellung geeignete Werkzeuge und Methoden auswählen, sich diese erschließen und zielgerichtet einsetzen</li> <li>• selbstständig aktuelle Entwicklungen des Themas verfolgen und im Projektalltag moderne Methoden einsetzen</li> <li>• das Zusammenspiel von Vorhersagedomäne, Datencharakteristik, Bewertungsmetrik und Geschäftsentscheidungen einschätzen</li> <li>• den Hintergrund und die Wertschöpfungskette analytischer Daten beschreiben</li> <li>• analytische Problemstellungen einordnen und Methoden sowie Werkzeuge lösungsorientiert einsetzen</li> <li>• die Tools des Python Data Science Stacks für die Problemlösung korrekt verwenden</li> <li>• den gesellschaftlichen Einfluss von Datensammlungs- und Auswertesystemen erkennen und eigene Tätigkeiten kritisch, gesamtgesellschaftlich reflektieren</li> </ul>
<b>Inhalte</b>	<ul style="list-style-type: none"> <li>• Methodisch kommen moderne Algorithmen der KI zum Einsatz für die Lösung von Clustering-, Regressions- und Klassifikationsproblemen</li> <li>• Umsetzungswerkzeug für die Implementierung ist die Programmiersprache Python mit dem dedizierten Data Science Stack</li> <li>• Verwendung von Python Bibliotheken für Datenanalyse und Umsetzung der KI-Architekturen in verteilten Systemen (bspw. Spark)</li> <li>• Besonderheiten von Streaming Systemen unter Aspekten des KI-Trainings</li> <li>• Aktuelle Methoden und Werkzeuge für die Analysekette von deskriptiver, diagnostischer, prädiktiver und präskriptiver Verarbeitung</li> <li>• Vorgehen um ausgehend von einer speziellen Datencharakteristik und einem konkreten Vorhersageziel Bewertungsmetriken und Algorithmik für die Problemlösung zu definieren</li> <li>• Methoden für Datenanonymisierung und diskriminierungsfreies Modelltraining</li> </ul>
<b>Vorleistungen und Modulprüfung</b>	<ul style="list-style-type: none"> <li>• Vorleistungen: <ul style="list-style-type: none"> <li>○ Keine</li> </ul> </li> <li>• Modulprüfung <ul style="list-style-type: none"> <li>○ Seminarfacharbeit zu einem vorgegebenen Analytics Problemfall mit abschließender Präsentation</li> </ul> </li> </ul>
<b>Literatur</b>	<ul style="list-style-type: none"> <li>• Jake VanderPlas, Python Data Science Handbook: Essential Tools for working with Data</li> <li>• Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython</li> </ul>

- 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