

Modulcode ^{1.}	Modulbezeichnung ^{2.}	Zuordnung ^{3.}
BAAI-4620 BAAI-8620	Data Integration/Data Mining (DIDM)	
	Studiengang ^{4.}	Bachelor Angewandte Informatik
	Fakultät ^{5.}	Gebäudetechnik und Informatik

Modulverantwortlich ^{6.}	Prof. Dr. Ines Rossak
Modulart ^{7.}	Pflichtmodul der Vertiefung Wirtschaftsinformatik sowie Wahlmodul für andere Vertiefungsrichtungen ab SPO2016/ Modul des Brückensemesters für das Masterstudium ab SPO2016
Angebotshäufigkeit ^{8.}	WS (Brückensemester) , SS (Pflicht-/Wahlmodul BA6)
Regelbelegung / Empf. Semester ^{9.}	Brückensemester für MA-AI ab SPO2016, BA6 für BA-AI ab SPO2016
Credits (ECTS) ^{10.}	5 CP
Leistungsnachweis ^{11.}	PL
Unterrichtssprache ^{12.}	Deutsch
Voraussetzungen für dieses Modul ^{13.}	Absolvierte Module Datenbanken 1 und Datenbanken 2
Modul ist Voraussetzung für ^{14.}	-
Moduldauer ^{15.}	1 Semester
Notwendige Anmeldung ^{16.}	Über CAJ, max. 24 Teilnehmer
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 Data Integration/ Data Mining	Rossak	S	24	1	4	60	65
Summe					4	60	65
Workload für das Modul ^{26.}						125	

Qualifikationsziele	<p>Die Studierenden können...</p> <ul style="list-style-type: none"> • Ziele, typische Einsatzbereiche und Szenarien für Data Integration und Data Mining benennen und erläutern • Begriff und verschiedene Prozessmodelle des Knowledge Discovery in Databases (KDD) verstehen und wiedergeben • Data Integration und Data Mining als Teilprozesse des KDD charakterisieren und die jeweils typischen Herausforderungen und Lösungsansätze verstehen, wiedergeben und auf ausgewählte Fallbeispiele anwenden • das CRISP-DM – Prozessmodell erläutern und selbständig auf ausgewählte Fallbeispiele anwenden • Grenzen und Gefahren von Data Mining erkennen und die Möglichkeiten des Datenmissbrauches, die sich aus Data Mining ergeben können, erfassen • für ausgewählte Fallbeispiele geeignete Data Mining Techniken sowie geeignete Algorithmen erkennen und einsetzen und die Auswahl begründet darstellen • die wichtigsten am Markt etablierten kommerziellen und nicht kommerziellen Tools benennen und mittels vorgegebener Kriterien evaluieren • sich mit geringer Anleitung selbständig in ein ausgewähltes Tool einarbeiten und dieses für ausgewählte Aufgabenstellungen einsetzen
Inhalte	<ul style="list-style-type: none"> • KDD – Begriff und Prozessmodelle mit Fokus auf CRISP-DM (detaillierte Behandlung der einzelnen Phasen) • Probleme, Methoden und Techniken der Datenintegration • Probleme, Methoden und Techniken der Datenexploration • Probleme, Methoden und Techniken der Datenanalyse <ul style="list-style-type: none"> ○ Gerichtete Analysetechniken (OLAP) ○ Ungerichtete Analysetechniken (Data Mining)
Vorleistungen und Modulprüfung	<p>Vorleistungen:</p> <ul style="list-style-type: none"> • keine <p>Modulprüfung:</p> <ul style="list-style-type: none"> • Projekt mit Präsentation der Meilensteine im Vorlesungszeitraum, Endpräsentation als mdl. Prüfung im PZR (Gruppenprüfung in Projektgruppen) • Meilensteine (10%), Endpräsentation (10%), Dokumentation 20%, Projekt an sich 40%
Literatur	<ul style="list-style-type: none"> • Rossak, I. (Hrsg.): Datenintegration, HANSER, 2013 • Ian H. Witten, Frank Eibe: Data Mining: Praktische Werkzeuge und Techniken für das maschinelle Lernen, Hanser 2001 • Helge Petersohn: Data Mining: Verfahren, Prozesse, Anwendungsarchitektur, Oldenbourg, 2005 • Roland Gabriel, Peter Gluchowski, Alexander Pastwa: Datawarehouse und Data Mining, W3I, 2009 • Matthew North: Data Mining for the Masses, 2012 • Markus Hofmann, Ralf Klinkenberg: RapidMiner: Data Mining Use Cases and Business Analytics Applications, Chapman&Hall, 2013 • Artikel aus BI-Spektrum, IS-Report, CIO-Magazin u.a. und white paper von BARC, TDWI und DGIQ • Handbücher der jeweiligen Tools sowie Online-Tutorials