



**Hochschule
Albstadt-Sigmaringen**
University of Applied Sciences

Berufsbegleitender Masterstudiengang

Data Science (M.Sc.)

Modulhandbuch

Stand: Sommersemester 2020



Inhalt

Vorbemerkung	2
2 Modulübersicht & Studienplan	3
3 Studiengangs-Kompetenzmatrix	4
4 Qualifikationsziel-Modul-Matrix.....	6
5 Modulbeschreibungen.....	8
10100 Programming for Data Science	8
10200 Mathematical Foundation for Data Science	9
10300 Data Mining	10
10400 Business Intelligence & Warehouse	12
20100 Databases.....	13
20200 WebData Integration	14
20300 Machine Learning.....	15
20400 Decision Support.....	17
30100 Big Data	18
30200 Advanced Machine Learning	20
30300 Text Mining	23
30400 Business Process and Big Data Use Cases	24
40100 Summer School.....	26
40200 Seminararbeit	27
50100 Advanced Statistics.....	28
50200 Web Mining.....	30
50300 Semantic Web Technologies	31
50400 Data Privacy & Data Compliance.....	33
60100 Masterthesis.....	35

Hinweis: Die Abkürzungen wurden aus der Studien- und Prüfungsordnung des Masterstudienganges Data Science entnommen und sind in § 36 der Studienprüfungsordnung Data Science erläutert.

Vorbemerkung

Der Studiengang ist als Fernstudium mit integriertem Blended-Learning-Ansatz modular mit Studienbriefen, Präsenz- und Online-Phasen sowie Betreuung durch Online-Tutoren und Professoren aufgebaut. Die Regelstudienzeit beträgt bis zum Erreichen des Master-Grades sechs Semester.

Das Studium vermittelt theoretische und praktische Kenntnisse in den Bereichen Datenanalyse, Datenextraktion und Dateninterpretation. Ziel des Studiums ist die Befähigung des Absolventen zu praktischen, konzeptionellen, wissenschaftlichen und wirtschaftlichen Tätigkeiten im Bereich Data Science.

Im 1. Semester werden die Grundlagen für Data Science in den Bereichen Business Intelligence, Data Warehouse, Data Mining sowie Programmiergrundlagen vermittelt. Aufbauend auf diesen Grundlagen werden im 2. bis 5. Semester die Kenntnisse mittels Modulen in den Bereichen Big Data, Machine Learning, Text- und Web Mining, Decision Support, Compliance, Business Process sowie Advanced Statistics vertieft.

Im 4. Semester wird ein zweiwöchiges Kompaktseminar durchgeführt, in welchem die Studierenden mit Unterstützung von Fachexperten aus der Industrie und Behörden Werkzeuge und Methoden der Datenanalyse kennenlernen. Darüber hinaus beginnen die Studierenden eine praktische Seminararbeit, die optimalerweise in Zusammenarbeit mit einem Unternehmen in der Anwendungsbranche erfolgt.

Mit der Masterthesis zeigen die Teilnehmer am Ende des Studiums, dass sie die Fähigkeiten besitzen, Theorie und Technik mit Reflexion auf die eigene berufliche Qualifizierung wissenschaftlich umzusetzen.

2 Modulübersicht & Studienplan

Die nachfolgende Abbildung zeigt die Verteilung der Module und ECTS-Punkte im Semesterverlauf:

	Fach- semester	Business Information	Data Analytics		Data Management
Vertiefungs- studium	6	Master-Thesis (25 ECTS) und Verteidigung (5 ECTS) Modul 60100, 30 ECTS			
	5	Data Privacy & Data Compliance Modul 50400, 5 ECTS	Semantic Web Technologies Modul 50300, 5 ECTS	Web Mining Modul 50200, 5 ECTS	Advanced Statistics Modul 50100, 5 ECTS
Praxisstudium	4	Summer School Modul 40100, 2,5 ECTS			
		Practical Work (Seminararbeit) Modul 40200, 7,5 ECTS			
Vertiefungs- studium	3	Business Process & Big Data Use Cases Modul 30400, 5 ECTS	Text Mining Modul 30300, 5 ECTS	Advanced Machine Learning Modul 30200, 5 ECTS	Big Data Modul 30100, 5 ECTS
	2	Decision Support Modul 20400, 5 ECTS	Machine Learning Modul 20300, 5 ECTS	Web Data Integration Modul 20200, 5 ECTS	Databases Modul 20100, 5 ECTS
Grundlagen- studium	1	Business Intelligence & Warehouses Modul 10400, 5 ECTS	Data Mining Modul 10300, 5 ECTS	Mathematical Foundations for Data Science Modul 10200, 5 ECTS	Programming for Data Science Modul 10100, 5 ECTS

Die Module einer thematischen Säule bauen inhaltlich auf einander auf. Die Teilnehmer werden kontinuierlich an schwierigere und komplexere Themen der Data Science herangeführt.

3 Studiengangs-Kompetenzmatrix

Kompetenzen	Fachkompetenz				
	Wissen		Fertigkeiten		
Ausprägung	Tiefe	Breite	Instrumentelle Fertigkeiten	systemische Fertigkeiten	Beurteilungsfähigkeit
Programming for Data Science		x	x	x	
Mathematical Foundations for Data Science	x			x	
Data Mining	x		x	x	x
Business Intelligence & Warehouse		x	x	x	
Databases	x		x	x	
WebData Integration	x		x	x	
Machine Learning	x		x	x	
Decision Support	x	x			x
Big Data	x		x	x	
Advanced Machine Learning	x		x	x	
Text Mining	x		x	x	
Business Process and Big Data Use Cases		x			x
Summer School		x			x
Seminararbeit	x		x	x	
Advanced Statistics	x		x	x	
Web Mining	x		x	x	
Semantic Web Technologies	x		x	x	
Data Privacy & Data Compliance		x			x
Master Thesis	x		x	x	

Kompetenzen	Personale Kompetenz					
	Sozialkompetenz			Selbstständigkeit		
Ausprägung	Team- /Führungs- fähigkeit	Mitgestaltung	Kommunikation	Eigenständigkeit/ Verantwortung	Reflexivität	Lernkompetenz
Programming for Data Science						
Mathematical Foundations for Data Science						
Data Mining						
Business Intelligence & Warehouse						
Databases						
WebData Integration						
Machine Learning						
Decision Support						
Big Data						
Advanced Machine Learning						
Text Mining						
Business Process and Big Data Use Cases	x	x	x			
Summer School	x	x	x			
Seminararbeit				x	x	x
Advanced Statistics						
Web Mining						
Semantic Web Technologies						
Data Privacy & Data Compliance	x	x	x			
Master Thesis				x	x	x

4 Qualifikationsziel-Modul-Matrix

Modul-Nr.	Qualifikationsziel (QuZ)	Summe der	Konzeptentwicklung	DS Problemstellungen erkennen, analysieren und lösen	Forschung und Entwicklung im DS Umfeld
	Modulbezeichnung				
10100	Programming for Data Science	5	1	2	2
10200	Mathematical Foundations for Data Science	5	1	2	2
10300	Data Mining	5	1	2	2
10400	Business Intelligence & Warehouse	5	2	2	1
20100	Databases	5	1	2	2
20200	WebData Integration	5	1	2	2
20300	Machine Learning	5	1	2	2
20400	Decision Support	5	1	2	2
30100	Big Data	5	1	2	2
30200	Advanced Machine Learning	5	1	2	2
30300	Text Mining	5	1	2	2
30400	Business Process and Big Data Use Cases	5	2	2	1

40100	Summer School	5	2	2	1
40200	Seminararbeit	5	2	2	1
50100	Advanced Statistics	5	1	2	2
50200	Web Mining	5	1	2	2
50300	Semantic Web Technologies	5	1	2	2
50400	Data Privacy & Data Compliance	5	2	2	1
60100	Master Thesis	5	2	2	1

5 Modulbeschreibungen

10100 Programming for Data Science

Modul: Programming for Data Science						
10100	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	1	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenzzeit	Selbststudium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen:					
	Kompetenz Wissen Die Studierenden sind in der Lage in R und Python funktional und datenorientiert zu programmieren und mit Data Frames umzugehen. <i>[Wissen, 7]</i>					
	Kompetenz Fertigkeiten Die Studierenden können den Stoff praktisch in der Programmiersprache R und Python für Analysen umsetzen <i>[Instrumentelle Fertigkeiten, 7]</i>					
	Sozialkompetenz Die Studierenden können statistische Sachverhalte anderen vermitteln. <i>[Kommunikation, 6]</i>					
	Selbstständigkeit Die Studierenden können selbstständig Analysen mittels der Programmiersprachen R und Python durchführen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	<ul style="list-style-type: none"> • Grundlagen Python: Schleifen, Variablen, grundlegende Daten-strukturen wie Listen, Tupel, Strings und Dictionaries. • Funktionale Programmierung mit Python: Map, Filter, Reduce, List-Comprehensions • Grundlagen Numpy • Grundpagen Pandas, Data Frames • Grundlagen R: Vektorisierte Operationen, Vektoren, Listen, Matrizen, Data Frames • Grundlagen der Visualisierung in Python und R mit Matplotlib und ggplot. 					
	Empfohlene Literaturangaben:					
	<ul style="list-style-type: none"> • Häberlein, T. (2017). Informatik: Eine Einführung mit Bash und Python. De Gruyter Oldenbourg-Verlag. • McKinney, W. (2012). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media, Inc. • Wollschläger, D. (2013). R Kompakt: Der Schnelle Einstieg in die Datenanalyse (Springer-Lehrbuch). Springer-Verlag • Chang, W. (2012). R Graphics Cookbook: Practical Recipes for Visualizing Data. O'Reilly Media, Inc. 					

4	Teilnahmevoraussetzungen: keine
5	Prüfungsformen: Klausur K60
6	Voraussetzungen für die Vergabe von Kreditpunkten: Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.
8	Modulverantwortliche(r): Prof. Dr. Häberlein Dozenten: Prof. Dr. Häberlein

10200 Mathematical Foundation for Data Science

Modul: Mathematical Foundation for Data Science						
10200	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	1	1 Semester	WS	
1	Lehrveranstaltung(en) Onlinevorlesungen, Präsenzwochenende, Fernstudien		Sprache Deutsch/ Englisch	Präsenz-zeit 20h	Selbst-studium 130h	Credits (ECTS) 5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden haben sich Anwendungsnahes Wissen in Stochastik, Statistik und Kombinatorik erworben und sind in der Lage dieses Wissen mit Hilfe der Programmiersprache R anzuwenden und in den nachfolgenden Veranstaltungen darauf aufzubauen. <i>[Wissen, 7]</i> <hr/> <i>Kompetenz Fertigkeiten</i> Nicht relevant <hr/> <i>Sozialkompetenz</i> Nicht relevant <hr/> <i>Selbstständigkeit</i> Die Studierenden erarbeiten sich den Inhalt selbstständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: <ul style="list-style-type: none"> • Grundbegriffe der Stochastik (Ereignis, Wahrscheinlichkeit, Bedingte Wahrscheinlichkeit, Unabhängigkeit, Zufallsvariablen) • Maßzahlen (Erwartungswert, Varianz, Median, usw.) • Grundlagen der Kombinatorik: Urnenmodell 					

	<ul style="list-style-type: none"> • Verteilungen, Dichtefunktionen, Zentraler Grenzwertsatz • Induktive Statistik, Likelihood, Regression, • Bayessche Statistik • Entropie, Entscheidungsbäume, Huffman-Codes <p><i>Empfohlene Literaturangaben:</i></p> <ul style="list-style-type: none"> • James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R (Springer Texts in Statistics). Springer-Verlag. • Lavine, M. (2007). Introduction to Statistical Thought. (online available as pdf) • Kerns, G. J. (2011). Introduction to Probability and Statistics Using R. (online available as pdf).
4	<p>Teilnahmevoraussetzungen:</p> <p>Grundlegendes Mathematisches Verständnis in Analysis (Ableitungen, Integration, Gleichungsumformung, usw).</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Häberlein Dozenten: Prof. Dr. Häberlein</p>

10300 Data Mining

Modul: Data Mining						
10300	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	1	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenzzeit	Selbststudium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen:					
	<p><i>Kompetenz Wissen</i></p> <p>Students will acquire fundamental knowledge of the techniques, opportunities and applications of data mining. Successful participants will be able to identify opportunities for applying data mining in an enterprise environment, select and apply appropriate techniques, and interpret the results. <i>[Wissen, 7]</i></p>					
	<p><i>Kompetenz Fertigkeiten</i></p> <p>Students learn to apply data mining techniques in business scenarios using state of the art data mining tools. <i>[Instrumentelle Fertigkeiten, 7]</i></p>					
	<p><i>Sozialkompetenz</i></p>					

	<p>Students learn to work as a team in order to solve a data mining project (case study). <i>[Team-/Führungsfähigkeit, 7]</i></p> <hr/> <p><i>Selbstständigkeit</i></p> <p>Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. In einer Projektarbeit während des Präsenzwochenendes lernen die Studierenden in kleineren Teams selbständig zu arbeiten. 7]</p>
3	<p>Inhalte:</p> <p>The course provides an introduction to advanced data analysis techniques as a basis for analyzing business data and providing input for decision support systems. The course covers the following topics:</p> <ul style="list-style-type: none"> • Goals and Principles of Data Mining • The Data Mining Process (Data Representation and Preprocessing) • Clustering (k-Means Clustering, Hierarchical Clustering, Density-based Clustering, Proximity Measures) • Classification (k-Nearest-Neighbors, Naïve Bayes, Decision Trees, Rule Induction, Support Vector Machines, Neural Networks, Model Evaluation, The Overfitting Problem, Parameter Optimization) • Association Analysis (Frequent Itemset Generation, Rule Generation, Interestingness Measures, Sequential Patterns) <hr/> <p><i>Empfohlene Literaturangaben:</i></p> <p>• Tan, P.-N., Steinback, M., Kumar, V. (2019). Introduction to Data Mining (2nd Edition). John Wiley & Sons</p>
4	<p>Teilnahmevoraussetzungen:</p> <p>Keine</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60 Projektarbeit Pj</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Bizer Dozenten: Prof. Dr. Bizer</p>

10400 Business Intelligence & Warehouse

Modul: Business Intelligence & Warehouse						
10400	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	1	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenzzeit	Selbststudium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden kennen den elementaren Aufbau von Data Warehouse Systemen und sind mit den zentralen Konzepten der Informationsvisualisierung vertraut <i>[Wissen, 7]</i> <hr/> <i>Kompetenz Fertigkeiten</i> Die Studierenden sind in der Lage, die Konzepte des Data Warehousing in realen Anwendungsszenarien konzeptionell und operativ umzusetzen. Sie können Kennzahlen mittels geeigneter Methoden visualisieren. <i>[Instrumentelle Fertigkeiten, 7]</i> <hr/> <i>Sozialkompetenz</i> Die Studierenden erarbeiten gemeinsam Lösungsansätze zu vorgegebenen Fragestellungen <i>[Mitgestaltung, 7]</i> <hr/> <i>Selbstständigkeit</i> Konzeption und Aufbau von Szenarien zur Erfassung, Aufbereitung und Analyse von Kennzahlen wird in Bezug auf die jeweiligen Anforderungen kritisch diskutiert. <i>[Reflexivität, 7]</i>					
3	Inhalte: Abgrenzung dispositive vs. operative Datenbestände (OLTP / OLAP) Data Warehouse Architekturen ETL-Prozesse (Datenqualität, Datenbereinigung, Transformation, etc.) Logische und semantische Datenmodelle für Data Warehouses (Star-/Snowflake-Schema,...) Implementierung von Data Warehouses (MOLAP, ROLAP,..) spezielle Aspekte im Zusammenhang mit Data Warehouses (slowly changing dimensions,...) Date Warehouse Projekte (agiles Vorgehen) Aufbau eines Data Warehousing Prozesses mittels SAP BW on HANA Visualisierung von Kennzahlen Konzeption und Aufbau von Dashboards Visual Analytics (Reduktion von Darstellungsdimensionen, Visualisierung von Objekten auf Grundlage von Unähnlichkeitsmaßen...) <hr/> <i>Empfohlene Literaturangaben:</i> Empfohlene Literaturangaben					

4	Teilnahmevoraussetzungen: Keine
5	Prüfungsformen: Klausur K60 Projektarbeit Pj
6	Voraussetzungen für die Vergabe von Kreditpunkten: Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.
8	Modulverantwortliche(r): Prof. Dr. Nemirovskij, Prof. Dr. Bernd Stauß Dozenten: Prof. Dr. Nemirovskij, Prof. Dr. Bernd Stauß

20100 Databases

Modul: Databases						
20100	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	2	1 Semester	SS	
1	Lehrveranstaltung(en) Onlinevorlesungen, Präsenzwochenende, Fernstudien		Sprache Deutsch/ Englisch	Präsenz-zeit 20h	Selbst-studium 130h	Credits (ECTS) 5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden - kennen Systeme und Techniken für die parallele Datenverarbeitung - kennen die Aufgabenstellungen aus dem Themengebiet von Big Data [<i>Wissen, 7</i>] <hr/> <i>Kompetenz Fertigkeiten</i> Lernergebnisse (Kompetenzen) bei: Die Studierenden - sind in der Lage die Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Databases zu erkennen, diese, basierend auf eigenem Wissen und durch die gezielte Recherche zu beschreiben, Lösungsansätze zu entwickeln und diese allein oder im Team umzusetzen. - sind in der Lage, eine anwendungsbezogene Evaluation von Daten, –Zugriffs- und –Verwaltungstechniken sowie von den diese Techniken implementierenden Systemen auszuführen, und darauf basierend eine zielgerechte Auswahl zu treffen. - sind in der Lage wissenschaftliche Beiträge im Themenbereich Databases eigenständig zu lesen und qualitative Vergleiche der gelesenen Beiträge systematisch zu präsentieren. [<i>Instrumentelle Fertigkeiten, 7</i>] <hr/> <i>Sozialkompetenz</i>					

	Nicht relevant 7]
	Selbstständigkeit Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i>
3	Inhalte: Datenbankentwurf Entity-Relationship-Modell Normalformen und Funktionale Abhängigkeiten Das relationale Modell SQL, Windows Functions Constraints Transaktionen, Stored Procedures, Trigger Indizes Relationale Algebra Python- und Javazugriff auf Datenbanken <i>Empfohlene Literaturangaben:</i> Alfons Kemper, Andre Eickler: Datenbanksysteme, De Gruyter Studium, 2015
4	Teilnahmevoraussetzungen: keine
5	Prüfungsformen: Klausur K60
6	Voraussetzungen für die Vergabe von Kreditpunkten: Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.
8	Modulverantwortliche(r): Prof. Dr. Eppler Dozenten: Prof. Dr. Eppler

20200 WebData Integration

Modul: WebData Integration						
20200	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	2	1 Semester	SS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen:					

	<p>Kompetenz Wissen</p> <p>Students will be able select and apply appropriate techniques for integrating and cleansing enterprise as well as Web data. Participants will acquire knowledge of the data integration process as well as the techniques that are used in each phase of the process. <i>[Wissen, 7]</i></p>
	<p>Kompetenz Fertigkeiten</p> <p>Students learn to apply data integration techniques in business scenarios <i>[Instrumentelle Fertigkeiten, 7]</i></p>
	<p>Sozialkompetenz</p> <p>Students learn to work as a team in order to succeed in a data integration project (case study) <i>[Team-/Führungsfähigkeit, 7]</i></p>
	<p>Selbstständigkeit</p> <p>Nicht relevant</p>
3	<p>Inhalte:</p> <ul style="list-style-type: none"> • Heterogeneity and Distributedness • The Data Integration Process • Web Data Formats • Schema Mapping and Data Translation • Identity Resolution • Data Quality Assessment • Data Fusion <p>Empfohlene Literaturangaben:</p> <p>AnHai Doan, Alon Halevy, Zachary Ives: Principles of Data Integration. Morgan Kaufmann, 2012. Luna Dong, Divesh Srivastava: Big Data Integration. Morgan & Claypool, 2015.</p>
4	<p>Teilnahmevoraussetzungen:</p> <p>Keine</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60 Projektarbeit Pj</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Bizer Dozenten: Prof. Dr. Bizer</p>

20300 Machine Learning

Modul: Machine Learning

20300		Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
		150 h	P	2	1 Semester	SS	
1	Lehrveranstaltung(en) Onlinevorlesungen, Präsenzwochenende, Fernstudien			Sprache Deutsch/ Englisch	Präsenz-zeit 20h	Selbst-studium 130h	Credits (ECTS) 5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden - kennen Grundbegriffe des Maschinellen Lernens - kennen Verfahren und Techniken für das Maschinelle Lernen - kennen die Aufgabenstellungen aus dem Themengebiet von Maschinellern Lernen <i>[Wissen, 7]</i> <i>Kompetenz Fertigkeiten</i> Lernergebnisse (Kompetenzen) bei: Die Studierenden - sind in der Lage die Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Machine Learning zu erkennen, diese, basierend auf eigenem Wissen und durch die gezielte Recherche zu beschreiben, Lösungsansätze zu entwickeln und diese allein oder im Team umzusetzen. - sind in der Lage, eine anwendungsbezogene Evaluation von Verfahren und Methoden des Maschinellen Lernens sowie von den diese Verfahren implementierenden Systemen auszuführen, und darauf basierend eine zielgerechte Auswahl zu treffen. - sind in der Lage wissenschaftliche Beiträge im Themenbereich Machine Learning eigenständig zu lesen und qualitative Vergleiche der gelesenen Beiträge systematisch zu präsentieren. <i>[Instrumentelle Fertigkeiten, 7]</i> <i>Sozialkompetenz</i> Die Studierenden können im Rahmen einer eigenständigen Arbeit neue Ansätze für eine Problemstellung im Bereich Machine Learning entwickeln und diese im Team umsetzen <i>[Team-/Führungsfähigkeit, 7]</i> <i>Selbstständigkeit</i> Die Studierenden sind in der Lage komplexe Aufgaben im Bereich Maschinelles Lernen verantwortungsvoll zu erfüllen, realistische Ziele zu definieren und diese konsequent zu verfolgen. <i>[Eigenständigkeit/Verantwortung, 7]</i>						
3	Inhalte: - Einführung und Grundbegriffe des Maschinellen Lernens - Lineare Modelle für Regression - Lineare Modelle für Klassifikation - Neuronale Netze und Backpropagation - Reinforcement Learning - Unüberwachte Lernverfahren - Implementierung/Anwendung ausgewählter Methoden mit Python, Numpy, Pandas, Scikit-learn <i>Empfohlene Literaturangaben:</i>						

	<ul style="list-style-type: none"> • Bishop, C. (2006). Pattern Recognition and Machine Learning. Springer. • Raschka, S. (2016). Python Machine Learning. Packt Publishing. • McKinney, W. (2013). Python for Data Analysis. O'Reilly.
4	Teilnahmevoraussetzungen: Keine
5	Prüfungsformen: Klausur K60
6	Voraussetzungen für die Vergabe von Kreditpunkten: Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.
8	Modulverantwortliche(r): Prof. Dr. Andreas Knoblauch Dozent. Prof. Dr. Andras Knoblauch

20400 Decision Support

Modul: Decision Support						
20400	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	2	1 Semester	SS	
1	Lehrveranstaltung(en)		Sprache	Präsenzzeit	Selbststudium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Students will acquire basic knowledge of the techniques, opportunities and applications of decision theory. Successful participants will be able to identify opportunities for decision support in an enterprise environment, select and apply appropriate techniques, and interpret the results. <i>[Wissen, 7]</i>					
	<i>Kompetenz Fertigkeiten</i> Nicht relevant					
	<i>Sozialkompetenz</i> Nicht relevant 7]					
	Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: The Module is based on the Textbook: Artificial Intelligence: A Modern Approach by Stuart Russell and Peter Norvig and will cover the following chapters: <ul style="list-style-type: none"> • Intelligent Agents • Logical Decision Making 					

	<ul style="list-style-type: none"> • Quantifying Uncertainty • Probabilistic Reasoning • Making Simple Decisions • Game Theory and Mechanism Design
	<p><i>Empfohlene Literaturangaben:</i></p> <ul style="list-style-type: none"> • Russell, S., Norvig, P. (2016). Artificial Intelligence: A Modern Approach (Third Global Edition). Prentice Hall.
4	<p>Teilnahmevoraussetzungen:</p> <p>Keine</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Stuckenschmidt Dozenten: Prof. Dr. Stuckenschmidt</p>

30100 Big Data

Modul: Big Data						
30100	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	3	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	<p>Lernergebnisse (learning outcomes), Kompetenzen:</p> <p><i>Kompetenz Wissen</i></p> <p>Die Studierenden</p> <ul style="list-style-type: none"> - kennen Systeme und Techniken für die parallele Datenverarbeitung - kennen die Aufgabenstellungen aus dem Themengebiet von Big Data [<i>Wissen, 7</i>] 					
	<p><i>Kompetenz Fertigkeiten</i></p> <p>Lernergebnisse (Kompetenzen) bei:</p> <p>Die Studierenden</p> <ul style="list-style-type: none"> - sind in der Lage die Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Big Data zu erkennen, diese, basierend auf eigenem Wissen und durch die gezielte Recherche zu beschreiben, Lösungsansätze zu entwickeln und diese allein oder im Team umzusetzen. 					

	<p>- sind in der Lage, eine anwendungsbezogene Evaluation von Daten, –Zugriffs- und – Verwaltungstechniken sowie von den diese Techniken implementierenden Systemen auszuführen, und darauf basierend eine zielgerechte Auswahl zu treffen.</p> <p>- sind in der Lage wissenschaftliche Beiträge im Themenbereich Big Data eigenständig zu lesen und qualitative Vergleiche der gelesenen Beiträge systematisch zu präsentieren. <i>[Instrumentelle Fertigkeiten, 7]</i></p> <hr/> <p>Sozialkompetenz</p> <p>Die Studierenden können im Rahmen einer eigenständigen Arbeit neue Ansätze für einen Big Data Prozess mit komplexer Aufgabenstellung entwickeln <i>[Team-/Führungsfähigkeit, 7]</i></p> <hr/> <p>Selbstständigkeit</p> <p>Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i></p>
3	<p>Inhalte:</p> <p>Vorlesung:</p> <ul style="list-style-type: none"> - Überblick zu NO-SQL-Datenbanken - Überblick zu Graphendatenbanken - Architekturen für verteiltes und paralleles Datenmanagement und Datenverteilung - Verteilte Anfragebearbeitung - Clustering, Map Reduce, YARN, Tez - Verteilte Datenbanken <ul style="list-style-type: none"> - Vertikale/horizontale Fragmentierung - Fragmentierungstransparenz - Transaktionskontrolle - Frameworks für Skalierung und Parallelisierung der Datenzugriffe am Beispiel von Apache Hadoop, Spark und verteilten RDBMS <p>Praktikum:</p> <p>Arbeiten mit Apache Hadoop, Spark Clustern, IBM Cloud, Azure, IBM Data Warehouse</p> <p>Arbeiten mit MongoDB, Apache Cassandra, Neo4J</p> <p>Arbeiten mit Injectiontools wie Apache Nifi, Talend, IBM NodeRed</p> <hr/> <p>Empfohlene Literaturangaben:</p> <p>Ramon Wartala: Hadoop: Zuverlässige, verteilte und skalierbare Big-Data-Anwendungen, Open Source Press</p> <p>Edward Capriolo, Dean Wampler, Jason Rutherglen: Programming Hive, O'Reilly</p> <p>Tom White: Hadoop. The definitive Guide, O' Reilly</p> <p>Tobias Trelle: MongoDB, Der praktische Einstieg</p> <p>Edward Capriolo, et. al: Programming Hive Erhard Rahm, et. al: Verteiltes und Paralleles Datenmanagement</p>
4	<p>Teilnahmevoraussetzungen:</p> <p>Kenntnisse von relationalen Datenbanken</p>
5	<p>Prüfungsformen:</p> <p>Klausur K 60</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p>

	Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.
8	Modulverantwortliche(r): Prof. Dr. Eppler Dozenten: Prof. Dr. Eppler

30200 Advanced Machine Learning

Modul: Advanced Machine Learning						
30200	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	3	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden - kennen wesentliche Begriffe des Maschinellen Lernens - kennen fortgeschrittene Verfahren und Techniken für das Maschinelle Lernen - kennen fortgeschrittene Aufgabenstellungen aus dem Themengebiet von Maschinellern Lernen Diese Vorlesungen vermitteln einen systematischen vereinheitlichenden Überblick über Methoden des maschinellen Lernens und deren Anwendungsmöglichkeiten. Nach Abschluss dieses Moduls soll der/die Studierende die wichtigsten Methoden kennen und verstehen, sowie in der Lage sein - je nach Problemstellung - geeignete Verfahren des Maschinellen Lernens auszuwählen, anzuwenden und zu evaluieren. <i>[Wissen, 7]</i> <hr/> <i>Kompetenz Fertigkeiten</i> Lernergebnisse (Kompetenzen) bei: Die Studierenden - sind in der Lage fortgeschrittene Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Machine Learning zu erkennen, diese, basierend auf eigenem Wissen und durch die gezielte Recherche zu beschreiben, Lösungsansätze zu entwickeln und diese allein oder im Team umzusetzen. - sind in der Lage, eine anwendungsbezogene Evaluation von fortgeschrittenen Verfahren und Methoden des Maschinellen Lernens sowie von den diese Verfahren implementierenden Systemen auszuführen, und darauf basierend eine zielgerechte Auswahl zu treffen. - sind in der Lage wissenschaftliche Beiträge im Themenbereich Machine Learning eigenständig zu lesen und qualitative Vergleiche der gelesenen Beiträge systematisch zu präsentieren. <i>[Instrumentelle Fertigkeiten, 7]</i> <hr/> <i>Sozialkompetenz</i> Die Studierenden können im Rahmen einer eigenständigen Arbeit neue Ansätze für eine Problemstellung im Bereich Machine Learning entwickeln und diese im Team umsetzen <i>[Team-/Führungsfähigkeit, 7]</i>					

	<p>Selbstständigkeit</p> <p>Die Studierenden sind in der Lage komplexe Aufgaben im Bereich Maschinelles Lernen verantwortungsvoll zu erfüllen, realistische Ziele zu definieren und diese konsequent zu verfolgen. <i>[Eigenständigkeit/Verantwortung, 7]</i></p>
3	<p>Inhalte:</p> <p>Aufbauend auf den Vorlesungen "Data Mining" und "Decision Support" vermittelt diese Vorlesung einen fundierten Überblick über Methoden des maschinellen Lernens und deren Anwendungen:</p> <ul style="list-style-type: none"> • Einfache lineare Modelle für Regression und Klassifikation • Neuronale Netzwerke • Kernel-Methoden & Support Vector Machines • Graphical Probabilistic Models • Kombination von Modellen • Lernen von Verhaltensmodellen • Implementierung/Anwendung ausgewählter Methoden mit Python, Numpy, Pandas, Scikit-learn - Convolutional Neural Networks und Deep Learning - Recurrent Neural Networks und LSTM - Kernel Methoden und Support Vector Machines - Graphical Probabilistic Models - Lernen Probabilistischer Modelle und Expectation Maximization - Modellkombination - Lernen von Verhaltensmodelle - Implementierung/Anwendung ausgewählter Methoden mit Python, Numpy, Scikit-learn, Tensorflow, Keras <hr/> <p>Empfohlene Literaturangaben:</p> <ul style="list-style-type: none"> - Bishop, C. (2006). Pattern Recognition and Machine Learning. Springer. - I. Goodfellow, Y. Bengio, A. Courville: Deep Learning. MIT Press, 2016. - F. Chollet: Deep Learning mit Python und Keras. MITP, 2018. - Raschka, S. (2016). Python Machine Learning. Packt Publishing. - McKinney, W. (2013). Python for Data Analysis. O'Reilly.
4	<p>Teilnahmevoraussetzungen:</p> <p>Keine</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Knoblauch Dozenten: Prof. Dr. Knoblauch</p>

30300 Text Mining

Modul: Text Mining						
30300	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	3	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Successful participants will be able to understand fundamental methods for Natural Language Processing, as well as being able to select, apply and evaluate the most appropriate techniques for a variety of different practical and application-oriented scenarios. [<i>Wissen, 7</i>] <hr/> <i>Kompetenz Fertigkeiten</i> Nicht relevant <hr/> <i>Sozialkompetenz</i> Nicht relevant 7] <hr/> Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. [<i>Eigenständigkeit/Verantwortung, 7</i>]					
3	Inhalte: In the digital age, techniques to automatically process textual content have become ubiquitous. Given the breakneck speed at which people produce and consume textual content online – e.g., on micro-blogging and other collaborative Web platforms like wikis, forums, etc. – there is an ever-increasing need for systems that automatically understand human language, answer natural language questions, translate text, and so on. This class will provide a complete introduction to principles and methods of Natural Language Processing (NLP). Covered topics include a complete introduction to all major sub-fields of NLP (syntax, semantics, etc.), as well as applications (e.g., Machine Translation, Information Extraction). <hr/> <i>Empfohlene Literaturangaben:</i> <ul style="list-style-type: none"> • Jurafsky, D., Martin, J. H. (2009). Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics (2nd ed.). Prentice-Hall. • Manning, C., Schütze, H. (1999). Foundations of Statistical Natural Language Processing. MIT Press. Cambridge, MA. 					
4	Teilnahmevoraussetzungen: Modul 10300 – Data Mining					
5	Prüfungsformen: Klausur K60 Projektarbeit Pj					

6	Voraussetzungen für die Vergabe von Kreditpunkten: Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.
8	Modulverantwortliche(r): Prof. Dr. Ponzetto Dozenten: Prof. Dr. Ponzetto, Prof. Dr. Glavaš

30400 Business Process and Big Data Use Cases

Modul: Business Process and Big Data Use Cases							
30400		Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
		150 h	P	3	1 Semester	WS	
1	Lehrveranstaltung(en) Onlinevorlesungen, Präsenzwochenende, Fernstudien			Sprache Deutsch/ Englisch	Präsenz- zeit 20h	Selbst- studium 130h	Credits (ECTS) 5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> In diesem Modul lernen die Studierenden den Anwendungsbezug von Big Data Analytics in der unternehmerischen Praxis. Das Modul wie Unternehmen mit Big Data Lösungen messbare Beiträge für die Wertschöpfung leisten. Hierzu lernen die Studierenden, wie Big Data Analytics Projekte gemanagt werden, welche Wertschöpfungs- und Businessmodelle in der Datenwirtschaft möglich sind und wie Big Data Technologien für smarte und zukunftsorientierte Geschäftsprozesse im Unternehmen genutzt werden können. <i>[Wissen, 7]</i> <i>Kompetenz Fertigkeiten</i> Die Studierenden können relevante und für Unternehmen wichtige betriebswirtschaftliche und technologische Anwendungsfälle der Big Data zu identifizieren, abzugrenzen und beschreiben Die Studierenden können die Anwendungsfälle in konkrete Projektstrukturen überführen und sind in der Lage, die erforderlichen Ressourcen hierfür zu identifizieren und zu evaluieren. Die Studierenden sind in der Lage, den kaufmännischen, prozessualen oder technologischen Wertbeitrag der Big Data Projekte zu evaluieren und zu dokumentieren. <i>[Instrumentelle Fertigkeiten, 7]</i> <i>Sozialkompetenz</i> Die Studierenden erlernen in Kleinteams praxisorientierte Aufgabenstellungen zu bearbeiten, zu präsentieren und zu verteidigen. <i>[Team-/Führungsfähigkeit, 7]</i> <i>Selbstständigkeit</i> Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i>						
3	Inhalte: • Geschäftsmodelle in der Datenwirtschaft • Kategorisierung und Vorstellung von praxisrelevanten Use cases • Vorgehensmodelle für BDA-Projekte • Management von BDA-Projekten						

	<ul style="list-style-type: none"> • Return on Investment von BDA-Projekten • Referenzarchitekturen für BDA-Systeme • Rollen und Ressourcen in BDA-Projekten • Aspekte der Data Governance und Data Compliance <hr/> <p><i>Empfohlene Literaturangaben:</i></p> <p>DeGraf, Rober: Managing your Data Science Project, apress, 2019 Marr, Bernard: Data Strategy, KoganPage, 2017 Stacey, Patrick: Managing Big Data Analytics Projects, Achamore Books, 2017 McGilvray, Danette: Executing Data Quality Projects, Elsevier Inc., 2008</p>
4	<p>Teilnahmevoraussetzungen:</p> <p>Keine</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60 Projektarbeit Pj</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Ruf Dozenten: Prof. Dr. Ruf</p>

40100 Summer School

Modul: Summer School						
40100	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	75 h	P	4	1 Semester	SS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Projektarbeit in Präsenz		Deutsch	75h	0h	2,5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden Die Studierenden lernen im Rahmen der SummerSchool(Kompakt-seminar) unter Anleitung von Experten Werkzeuge und Methoden zur Analyse kennen und geeignet anzuwenden. <i>[Wissen, 7]</i>					
	<i>Kompetenz Fertigkeiten</i> Lernergebnisse (Kompetenzen) bei: Die Studierenden sind in der Lage die Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Data Science zu erkennen, diese, basierend auf eigenem Wissen und durch die gezielte Recherche zu beschreiben, Lösungsansätze zu entwickeln und diese allein oder im Team umzusetzen.. <i>[Instrumentelle Fertigkeiten, 7]</i>					
	<i>Sozialkompetenz</i> Die Studierenden können im Rahmen einer Teamarbeit neue Ansätze für einen Data Scienceprozess mit komplexer Aufgabenstellung entwickeln <i>[Team-/Führungsfähigkeit, 7]</i>					
	<i>Selbstständigkeit</i> Die Studierenden sind in der Lage komplexe Aufgaben verantwortungsvoll zu erfüllen, realistische Ziele zu definieren und diese konsequent zu verfolgen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: Einführung in DataScience Tools und -werkzeuge (1 Woche) Projektarbeit an konkreten Datensätzen (1 Woche)					
4	Teilnahmevoraussetzungen: keine					
5	Prüfungsformen: Projektarbeit und Referat					
6	Voraussetzungen für die Vergabe von Kreditpunkten:					

	Während einer Präsenzwoche ist eine Gruppenarbeit zu erstellen und diese zu präsentieren.
7	Modulverantwortliche(r): Prof. Dr. Eppler Dozenten: Prof. Dr. Ruf, Prof. Dr. Herda, et.al.

40200 Seminararbeit

Modul: Seminararbeit						
40200	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	225 h	P	4	1 Semester	SS	
1	Lehrveranstaltung(en) Fernstudien, Hausarbeit		Sprache Deutsch/ Englisch	Präsenz-zeit 0h	Selbst-studium 225h	Credits (ECTS) 7,5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden Der Studierende ist in der Lage das in den Semestern 1 bis 3 erlernte Wissen in einer wissenschaftlichen Arbeit umzusetzen [Wissen, 7]					
	<i>Kompetenz Fertigkeiten</i> Lernergebnisse (Kompetenzen) bei: Die Studierenden - sind in der Lage die Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Data Science in einer wissenschaftlichen Arbeit umzusetzen und zu präsentieren [Instrumentelle Fertigkeiten, 7]					
	<i>Selbstständigkeit</i> Die Studierenden sind in der Lage komplexe Aufgaben verantwortungsvoll zu erfüllen, realistische Ziele zu definieren und diese konsequent zu verfolgen. [Eigenständigkeit/Verantwortung, 7]					
3	Inhalte: Alle Themengebiete aus dem Bereich Data Science					
4	Teilnahmevoraussetzungen: keine					
5	Prüfungsformen: Praktische Arbeit (5 ECTS) Referat (2,5 ECTS)					

6	Voraussetzungen für die Vergabe von Kreditpunkten: keine
7	Modulverantwortliche(r): Alle Dozenten des Studienganges Data Science

50100 Advanced Statistics

Modul: Advanced Statistics						
50100	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	5	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenzzeit	Selbststudium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden sind in der Lage erweitern statistische Techniken zur statistischen Analyse von Datensätzen anzuwenden. Die Studierenden verstehen die diesen Techniken zugrundeliegenden theoretischen Konzepte, sind in der Lage diese unter zu Hilfenahme von Python und/oder R-Bibliotheken anzuwenden und entwickeln ein grundlegendes Verständnis dafür, was moderne Bayessche Statistik, Methoden der multivariaten Statistik, der EM-Algorithmus und probabilistic Programming leisten können und für welche Anwendungen diese einsetzbar sind. <i>[Wissen, 7]</i> <i>Kompetenz Fertigkeiten</i> Die Studierenden können die theoretischen Inhalte praktisch mit verschiedenen Systemen und Tools umsetzen. <i>[Instrumentelle Fertigkeiten, 7]</i> <i>Sozialkompetenz</i> Die Studierenden können statistische Sachverhalte anderen vermitteln. <i>[Kommunikation, 6]</i> Die Studierenden erarbeiten sich den Inhalt selbstständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: <ul style="list-style-type: none"> • Multivariate Verteilungen • Darstellung Multivariater Verteilungen • Multivariate Regression und Multiple Discriminant Analyses • MANOVA • Gaußsche Mischmodelle • EM-Algorithmus • Bayessche Statistik • Markov Chains • Markov-Chain-Monte-Carlo Sampler (Gibbs, Metropolis-Hastings, No-U-Turn, usw) • Probabilistic-Programming-Bibliotheken 					

	<ul style="list-style-type: none"> • Theano, PyMC3, Stan • Probabilistic Modeling <hr/> <p><i>Empfohlene Literaturangaben:</i></p> <ul style="list-style-type: none"> • Davidson-Pilon, C. (2015). Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference. Addison Wesley. • Bruce, P., Bruce, A. (2017). Practical Statistics for Data Scientists. O'Reilly.
4	<p>Teilnahmevoraussetzungen:</p> <p>Modul 10100 – Programming for Data Science</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Häberlein Dozenten: Prof. Dr. Häberlein</p>

50200 Web Mining

Modul: Web Mining						
50200	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	5	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Students will acquire knowledge of the foundation, techniques and applications in the field of Web mining. <i>[Wissen, 7]</i> <hr/> <i>Kompetenz Fertigkeiten</i> Nicht relevant <hr/> <i>Sozialkompetenz</i> Nicht relevant <hr/> Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. In der Projektarbeit während des Präsenzwochenendes erarbeiten die Studierenden selbständig neue Inhalte. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: Structured and unstructured data available on the Web provide us with a goldmine of information that has the potential to enable cutting-edge intelligent applications. This class covers a variety of topics focused on mining techniques for Web data, including extracting knowledge from Web content (Web Content Mining), the link structure of the Web (Web Structure Mining), as well as mining usage data gathered by Web applications (Web Usage Mining). <hr/> <i>Empfohlene Literaturangaben:</i> <ul style="list-style-type: none"> • Liu, B. (2011). Web Data Mining (2nd Edition). Springer. • De Nooy, W., Mrvar, A., Batagelj, V. (2011). Exploratory Social Network Analysis with Pajek. Cambridge University Press. • Jannach, D. (2011). Recommender Systems: An Introduction. Cambridge University Press. 					
4	Teilnahmevoraussetzungen: Modul 10300 – Data Mining					
5	Prüfungsformen: Klausur K60 Projektarbeit Pj					
6	Voraussetzungen für die Vergabe von Kreditpunkten:					

	Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben und eine Projektarbeit in Gruppen anzufertigen.
8	Modulverantwortliche(r): Prof. Dr. Ponzetto Dozenten: Prof. Dr. Ponzetto, Prof. Dr. Glavaš

50300 Semantic Web Technologies

Modul: Semantic Web Technologies						
50300	Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
	150 h	P	5	1 Semester	WS	
1	Lehrveranstaltung(en)		Sprache	Präsenz-zeit	Selbst-studium	Credits (ECTS)
	Onlinevorlesungen, Präsenzwochenende, Fernstudien		Deutsch/ Englisch	20h	130h	5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> The participants of this course learn about principles and applications of Semantic Web standards. They become familiar with their technical foundations such as representation and query languages, or logical inference. After taking this course, the students will be aware of the problems and benefits of semantic technologies in the context of tasks such as knowledge management, information search and data integration, and they will be capable of judging the applicability of these technologies for addressing practical challenges. <i>[Wissen, 7]</i>					
	<i>Kompetenz Fertigkeiten</i> Nicht relevant					
	<i>Sozialkompetenz</i> Nicht relevant					
	Die Studierenden erarbeiten sich den Inhalt selbständig anhand von Studienbriefen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: This course gives an introduction to the technical foundations of Semantic Web Technologies, including knowledge representation and query languages, as well as logical inference. More specifically, it covers the following contents: <ul style="list-style-type: none"> • Vision and Principles of the Semantic Web • Representation Languages (XML, RDF, RDF Schema, OWL) • Knowledge Modeling: Ontologies and Linked Data • Logical Reasoning in RDF and OWL • Commercial and Open Source Tools and Systems 					
	<i>Empfohlene Literaturangaben:</i>					

	• Hitzler, P., Krötzsch, M., Rudolph, S. (2009). Foundations of Semantic Web Technologies. Chapman & Hall/CRC.
4	Teilnahmevoraussetzungen: Keine
5	Prüfungsformen: Klausur K60
6	Voraussetzungen für die Vergabe von Kreditpunkten: Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.
8	Modulverantwortliche(r): Prof. Dr. Stuckenschmidt Dozenten: Prof. Dr. Stuckenschmidt

50400 Data Privacy & Data Compliance

Modul: Data Privacy & Data Compliance							
50400		Workload	Modulart	Studiensemester	Dauer	Häufigkeit	
		150 h	P	5	1 Semester	WS	
1	Lehrveranstaltung(en) Onlinevorlesungen, Präsenzwochenende, Fernstudien			Sprache Deutsch/ Englisch	Präsenz-zeit 20h	Selbst-studium 130h	Credits (ECTS) 5
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Nach Abschluss des Moduls kennen die Studierenden die wichtigsten US-, EU- und deutschen Datenschutznormen und sind in der Lage, Unterschiede herauszustellen. Sie kennen die Voraussetzungen der Datenverarbeitung und können sie rechtlich würdigen sowie in den betrieblich-organisatorischen Kontext einordnen. Die Studierenden haben auf der Grundlage von Praxisfällen gelernt, wie Projekte aus Datenschutzperspektive zu bearbeiten sind. Sie kennen die Schutzziele von Europäischer Datenschutz-Grundverordnung als auch weiteren Datenschutzbestimmungen und können deren Berücksichtigung dokumentieren und bewerten. Die Studierenden kennen technische Lösungen zur Umsetzung des Datenschutzes und sind befähigt, diese auf konkrete Use Cases anzuwenden. <i>[Wissen, 7]</i> <i>Kompetenz Fertigkeiten</i> Die Studierenden sind in der Lage - datenschutzrechtlich relevante Situationen zu erkennen, und haben die Fähigkeit, zu entscheiden, welche Gesetzesmaterie sie heranziehen müssen, um die datenschutzrechtliche Bestimmungen für einen konkreten Anwendungsfall zu ermitteln. - sachgebietsbezogene Fragen der Gestaltung und Steuerung von berufspraktischen Kontexten unter Berücksichtigung technischer, organisatorischer und rechtlicher Faktoren zu reflektieren und mündlich oder schriftlich zu erörtern. - die Relevanz von wissenschaftlichem und historischem Wissen für den eigenen berufspraktischen Kontext einzuschätzen und wissenschaftliche Fachliteratur zu recherchieren und reflektieren. - sich differenziert und allgemeinverständlich schriftlich und mündlich zu Fragestellungen in datenschutzrechtlichen Kontexten äußern. - ihr Wissen und Verständnis sowie ihre Fähigkeiten zur Problemlösung auch in neuen und unvertrauten Situationen anzuwenden, die in einem breiteren oder multidisziplinären Zusammenhang mit dem Datenschutz stehen. <i>[Instrumentelle Fertigkeiten, 7]</i> <i>Sozialkompetenz</i> Die Studierenden sind in der Lage effektiv als Einzelner oder in Teams zu arbeiten, gepaart mit der Fähigkeit, lebensbegleitend zu lernen. Sie können effektiv mit der Wissensgemeinschaft der Datenschützer und den mit der Datenschutzumsetzung betroffenen Stakeholdern kommunizieren. <i>[7]</i> <i>Selbstständigkeit</i> Die Studierenden können aktive und individuell nützliche Ziele und Prioritäten setzen und diese verfolgen. Neben der Beschäftigung mit kurzfristigen Lösungen und nützlichen Methoden werden Selbstreflexionsprozesse angeregt, die sich mittel- und langfristig positiv auf die Entwicklung von Zeit-						

	und Selbstmanagement und Leistungsfähigkeit im Berufsleben auswirken. [Eigenständigkeit/Verantwortung, 7]
3	<p>Inhalte:</p> <ul style="list-style-type: none"> • Grundlagen des Datenschutzes • Abgrenzung zu anderen Grundrechten und Rechten • Normativer Rahmen für Data Privacy & Data Compliance • Big-Data-Anwendungen & datenschutzrechtliche Fragestellungen • Technische Maßnahmen zur Gewährleistung des Datenschutzes • Auftrags(daten)verarbeitung • Anforderungen an den betrieblichen Datenschutzbeauftragten • Aufbau einer Datenschutzorganisation <hr/> <p><i>Empfohlene Literaturangaben:</i></p> <ul style="list-style-type: none"> • Rüpke, Giselher/v. Lewinski, Kai/Eckardt, Jens (2018). Datenschutzrecht, Grundlagen und Neugestaltung. München. • Mayer-Schönberger, Viktor/Cukier, Kenneth (2017). Big Data. London • Koch, F. A. (2015). Big Data und der Schutz der Daten. ITRB, 13-20. • Weimer, L. (Hrsg.), (2017). Datenschutz, IT-Sicherheit & Cyber-Risiken (4. Auflage). Heidelberg.
4	<p>Teilnahmevoraussetzungen:</p> <p>Keine</p>
5	<p>Prüfungsformen:</p> <p>Klausur K60</p>
6	<p>Voraussetzungen für die Vergabe von Kreditpunkten:</p> <p>Am Präsenzwochenende ist eine Klausur (60 min) zu schreiben.</p>
8	<p>Modulverantwortliche(r):</p> <p>Prof. Dr. Ruf Dozenten: Prof. Dr. Ruf, Frau Zierau</p>

60100 Masterthesis

Modul: Masterthesis						
60100	Workload 900 h	Modulart P	Studiensemester 6	Dauer 1 Semester	Häufigkeit SS	
1	Lehrveranstaltung(en) Thesis, Verteidigung		Sprache Deutsch/ Englisch	Präsenz-zeit 0h	Selbst-studium 900h	Credits (ECTS) 30
2	Lernergebnisse (learning outcomes), Kompetenzen: <i>Kompetenz Wissen</i> Die Studierenden Mit der Master-Arbeit erbringt der Student den Nachweis, dass er unter Anleitung selbstständige umfangreiche wissenschaftliche Themen bearbeiten kann. Er zeigt damit, dass er in der Lage ist, praxisorientierte oder theoretische Themenstellungen nach wissenschaftlichen Kriterien zu analysieren, zu strukturieren und ergebnisorientiert zu bearbeiten. Die Master-Thesis dokumentiert seine Arbeit und erfüllt die Kriterien eines wissenschaftlichen Berichts. Der Student erläutert und begründet seine Vorgehensweisen, Methoden und seinen Lösungsweg. Mit der Verteidigung wird das erworbene Wissen des Studenten im Zusammenhang geprüft. Er zeigt, dass er das im Studium erworbene Wissen zur Lösung umfassender Probleme der Data Science anwenden kann. <i>[Wissen, 7]</i> <hr/> <i>Kompetenz Fertigkeiten</i> Lernergebnisse (Kompetenzen) bei: Die Studierenden - sind in der Lage die Problem- und Aufgabenstellungen mit Bezug auf das Themengebiet Data Science in einer grossen wissenschaftlichen Arbeit umzusetzen und zu präsentieren <i>[Instrumentelle Fertigkeiten, 7]</i> <hr/> <i>Selbstständigkeit</i> Die Studierenden sind in der Lage komplexe Aufgaben verantwortungsvoll zu erfüllen, realistische Ziele zu definieren und diese konsequent zu verfolgen. <i>[Eigenständigkeit/Verantwortung, 7]</i>					
3	Inhalte: Alle Themengebiete aus dem Bereich Data Science					
4	Teilnahmevoraussetzungen: keine					
5	Prüfungsformen: Thesis 25 ECTS Verteidigung M40 5 ECTS					
6	Voraussetzungen für die Vergabe von Kreditpunkten:					

	Bestehen die Masterthesis (schriftliche Ausarbeitung). Bestehen die mündliche Prüfung/Verteidigung
7	Modulverantwortliche(r): Alle Dozenten des Studienganges Data Science