Module Guide

M. Sc. non-consecutive full-time program
Big Data & Business Analytics

School of Information, Media and Design

DAS CORE PRINZIP

HOCHSCHULE SRH HEIDELBERG
Intelligence in Learning

STAATLICH ANERKANNTE HOCHSCHULE

Table of Contents

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3035 – First Steps into Case Studies

Semester	Annually offered	Duration (weeks)	Type	ECTS Points	Student's workload (1 ECTS appropriates 25 h; for exceptions view appendix 2b of the examination regulations)		
1	2x	5	Compulsory module	6	- 150 hours, thereof: - 30 in class, - 105 private studies, - 15 coaching		

Requirements for enrolment	Applicability	Type(s) of Exam	Teaching and learning methods	Module coordinator
None	In all following modules	- Project work and development - Learning diary	- Data research,- analysis and preparation - Case studies - Team work - Problem-oriented learning - Exercises	Prof. Dr. Ajinkya Prabhune

Learning objectives

<u>Abstract</u>

The main function of this module is to lay the essential knowledge foundation for all later modules. Students analyze various practical examples of Big Data projects and as teams they perform well-defined Big Data case studies which involve the whole process of a Big Data project: Definition of a discrete problem within a business – Data acquisition and cleansing – Data saving – Data analysis and interpretation – Data visualization and communication of the results provided by Data analytics - Recommendation of further actions. This setup enables the Big Data project being embedded into a business context. Thus, Students are enabled to interconnect commercial necessities and decision making with ethical issues during the Big Data project, which vice versa avoids an approach exclusively catering technical as well as analytical requirements. The datasets

used in the case studies are prepared by the SRH University and distributed to the students using a cloud platform. The concepts, methods and tools learned in this module will be repeated and intensified in the following modules – this especially affects the modules Case Studies I and II. Parallel to the Case Studies the students are taught basic scientific competences: They understand the essential aspects of scientific work and are able to plan and structure a scientific process. Furthermore, they show profound knowledge in the application of tools and methods during the scientific writing process.

Learning outcomes Specialist Competence

| The students know the fundamental aspects of Big Data science.
| They are able to define the "five Vs" of Big Data (Volume, Velocity, Variety, Value and Veracity) as well as the different phases of a Big Data project. They may explain them in well-defined practical sessions in an application-oriented manner.
| They know diverse practical examples of Big Data projects and are able to explain their approach in business context as well as to compare with each other.
| They understand the different phases of a Big Data project and may explain them in context with Big Data projects.
| They are able to structure their scientific work as well as their results
| They gain results based on appropriate scientific criteria, e.g. objectivity, validity and reliability.

Learning outcomes Method Competence

| The students are able to understand various types of project management and may apply as well as validate their knowledge in teams. Moreover, they understand the connection between business offers and the Big Data Life Cycle.

Learning outcomes Social Competence:

| The students know the fundamental tools to organize working processes in virtual teams as well approaches dealing with team-intern conflicts. They are able to use both tools to accomplish results mutually.

Learning outcomes Personal Competence

| The students are able to take their role within the virtual team parallel to their occupational activities and organize multiple tasks (i.e. occupation, private life and studies) simultaneously.

| Course contents |

Course content catering the aforementioned competences

| Five Vs: Volume, Velocity, Variety, Value and Veracity
| Big Data Life Cycle: Generation and collection of data, Data processing and storage, Data analysis, publishing, archiving and recall.
| Best Practices
| Cloud Computing
| Spectrum of Big Data solutions on the market: Hadoop / Map Reduce / Spark / AWS / Google / R or SAS / Lumira or Tableau
| Organization and management of Big Data projects
| Principles of scientific work
    o Scientific quality criteria
    o Scientific Methodology
    o Criteria to evaluate scientific works
    o Research, classification and evaluation of scientific literature
    o Scientific writing
    o Lead and host academic discussions

Recommended literature for preparation and follow-up

| Cielen D & Meysman A: Introducing Data Science, Manning Verlag, 2016

| Gupta D: Applied Analytics through Case Studies Using SAS & R, 1st edition, Apress, 2018.

| Marr B: Big Data in Practice, 1st edition, Wiley, 2016.

| Kromer P & Russell Jurney: Big Data for Chimps: A Guide to Massive-Scale Data Processing in Practice

| Wickham H: R for Data Science, 1st edition, O'Reilly, 2017.

Scientific work

| Heard SB: The Scientist's Guide to Writing: How to Write More Easily and Effectively throughout Your Scientific Career, Princeton University Press, 2016.

| Skern T: Writing Scientific English: A workbook, 2nd edition, UTB, 2011.

Links

| Google Scholar
| DBLP
| IEEE Computer Society
| IEEE TVCG camera ready document guidelines

Literature management software

| Citavi

Constructive Alignment
Using the examination types Project work, development and learn journal the students are given the opportunity to reflect and document practically their progress in learning as well as their scientific abilities. Through the analysis of practical examples and the performance of a complete case study using the business perspective as well as the multiple forms of Data handling (Choice, evaluation, cleansing, providing, analysis and communication) the students get a first glimpse of the technical, organizational and methodological principals of Big Data and are also able to interconnect them directly with the different aspects and phase of a Big Data project. Another aspect for the choice of the aforementioned examination types is that they enable a step-by-step improvement in skills and fit optimally to the practice oriented character of this module. Virtual teams are formed to enable the studiablity parallel to the students' main occupation. These teams are provided the main course material (data, software, scripts, literature) via a cloud platform and they may also store and share their progresses. Additional coaching of the teams during the module is provided via live chats and e-learning.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3036 – Case Studies 1

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
2	2x	20	Compulsory module	8	- 200 hours, thereof: - 20 in class, - 140 private studies, - 40 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
- First Steps Into Case Studies - Data Engineering - Analytics II	In all following modules	- Project work and development - Presentation	- Data research,- analysis and preparation - Case study - Working in teams - Exercises	Prof. Dr. Ajinkya Prabhune

Learning objectives

Abstract:

Similar to the module "First Steps Into Case Studies", a Big Data project is performed during Case Studies 1, with the difference that the focus lays on specific company compartments: Marketing, production, logistics and sales management. The main methodology used in "Case Studies 1" originates from data mining. In the beginning of the module the students are given the task to develop a Big Data question out of a specific company problem. It is a key element of Big Data projects to pose good and adequate questions; this crucial step is maintained by creativity techniques. Subsequently, a Big Data project is planned and executed to solve this question. Finally, the results of this analysis are visualized and used to develop further guidance.

This project is accompanied by the modules "Data Storytelling and communication" and "Privacy, Ethics and International Law".

The case studies base on the competencies and skills gained from the modules "First Steps Into Case Studies" and "Data Engineering". Basic knowledge in the fields "Data Management" and "Design Principles" are recommended but not required. Additional theoretical input in Data visualization originates from the module "Data Storytelling and Communication". It is no prerequisite to successfully accomplish the module "Data Storytelling and Communication" to complete "Case Studies 1".

As in "First Steps Into Case Studies", the case studies are performed within virtual teams, who interact with each other via cloud solutions. Each team is coached by the professors using live chat and synchronous E-learning.

Learning outcomes Specialist Competence:

| The students are able to transfer a company problem into a Big Data question as well as planning and performing it afterwards.
| They identify the data being necessary for this question and are able to estimate properly the value of the data in context of the problem.
| They may prepare data for Data Mining.
| They execute a data mining analysis with the help of established tools and software.
| They are capable to adequately visualize and communicate the results in context of the developed problem

Learning outcomes Method Competence:

| The students are able to specifically apply creativity techniques to develop a problem and to identify

required data.

| They evaluate properly the applicability of methods and tools for the different phases of the Big Data project in context of a certain project and are able to select and execute the adequate methods.

| They interpret and evaluate the results of the analytics process with regard to the developed Big Data problem

Learning outcomes Social Competence:

| The students know the fundamental tools to organize working processes in virtual teams as well approaches dealing with team-intern conflicts. They are able to use both tools to accomplish results mutually.

Learning outcomes Personal Competence:

| The students are able to take their role within the virtual team parallel to their occupational activities and organize multiple tasks (i.e. occupation, private life and studies) simultaneously.

Course content

Course content catering the aforementioned competences

| Project management
| Organization and management
| Creativity techniques and formulations of problems
| Big Data architectures
| Data Mining / Text Mining
| Storage and Retrieval Tools
| Data Mining Tools, methods and techniques

Recommended literature for preparation and follow-up:

| Buttcher S *et al.*: Information Retrieval: Implementing and Evaluating Search Engines, Reprint, MIT Press, 2016

| Gupta D: Applied Analytics through Case Studies Using SAS & R, 1st edition, Apress, 2018

| Han J *et al.*: Data Mining: Concepts and Techniques, Elsevier/Morgan Kaufmann, Amsterdam, 2006

| Kantardzic M: Data Mining, Wiley, 2011

| Koster K: International Project Management, Sage Publications Ltd., 2009

| Liu B: Web Data Mining: Exploring Hyperlinks, Contents and Usage Data (Data-Centric Systems and Applications), 2. Auflage, Springer, 2011

| Marr B: Big Data in Practice, 1st edition, Wiley, 2016

| Marsland S: Machine Learning - An Algorithmic Perspective, 2nd new edition, Taylor & Francis, 2014

| Moitra A: Algorithmic Aspects of Machine Learning, Reprint, Cambridge University Press, 2018

| Nussbaumer Knaflic C: Storytelling with Data: A Data Visualization Guide for Business Professionals, 1st edition, Wiley, 2016

| Runkler TA: Data Analytics – Models and Algorithms for Intelligent Data Analysis, 2nd edition, Springer Vieweg, 2016

| Verhoef P et al.: Creating Value with Big Data Analytics, Routledge, 2016

| Witten IH *et al.*: Data Mining: Practical Machine Learning, Tools and Techniques, 4th edition, Morgan Kaufmann, 2016

| Wickham H: R for Data Science, 1st edition, O'Reilly, 2017

Constructive alignment

The examination forms project work and presentation evaluates the students' abilities to application-specifically document their learning progress. Through the application of the skills and competencies acquired in the previous modules during "case studies 1" the students are given the opportunity to intensify their knowledge in a holistic manner. Furthermore, the students learn to evaluate the applicability and cooperation of methods, techniques and tools in a context of a certain project. The project work is an adequate examination for project module because of its emphasis on the visualization and communication of the results/recommendations.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3037 – Case Studies 2

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)	
3	2x	20	Compulsory module	8	- 200 hours, thereof: - 20 in class, - 140 private studies, - 40 coaching	

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
- First Steps In Case Studies - Data Engineering - Analytics II	In all following modules	- Project work and development - Presentation	- Data research, analysis and processing - Project work - Team work - Learning diary	Prof. Dr. Ajinkya Prabhune

Learning objectives

Abstract

Similar to the case studies in the modules "First Steps Into Case Studies" and "Case Studies 1", the module "Case studies 2" includes the generation of a Big Data problem which lays the foundation for a Big Data project. The results gained from this project are visualized and communicated as a business-oriented advice.

The main focus of the Big Data projects performed in "Case Studies 2" lays on predictive analytics. Additional theoretical input is provided by the simultaneously occurring modules "Data Storytelling and Communication", "Data Management: Data curation and modelling" and "Analytics III: Predictive Analytics". The results acquired from the former two case study modules are also implemented in "Case Studies 2". Fundamental knowledge in "Design Principals" is favorable, but not a prerequisite. The successful accomplishment of "Data Storytelling and Communication" is not required to complete "Case studies 2". As in the case studies of the previous two semesters the Big Data projects are performed in virtual teams which are interconnected via cloud solutions. Additionally, the groups are coached by the teaching staff via live chat and synchronous E-Learning

Learning outcomes Specialist Competence:

- The students are able to transfer a company problem into a Big Data question as well as planning and performing it afterwards.
- They identify the data being necessary for this question and are able to estimate properly the value of the data in context of the problem.
- They may prepare data for Predictive analytics.
- They execute a predictive analysis with the help of established tools and software.
- They are capable to adequately visualize and communicate the results in context of the developed problem

Learning outcomes Method Competence:

- The students are able to specifically apply creativity techniques to develop a problem and to identify required data.
- They evaluate properly the applicability of methods and tools for the different phases of the Big Data project in context of a certain project and are able to select and execute the adequate methods.
- They interpret and evaluate the results of the analytics process with regard to the developed Big Data problem

Learning outcomes Social Competence:

- The students know the fundamental tools to organize working processes in virtual teams as well approaches dealing with team-intern conflicts. They are able to use both tools to accomplish results mutually.

Learning outcomes Personal Competence:

- The students are able to take their role within the virtual team parallel to their occupational activities and organize multiple tasks (i.e. occupation, private life and studies) simultaneously.

Course content

11

Course content catering the aforementioned competences

| Project management

| Organization and management

| Creativity techniques, formulation of questions

| Data management

| Big Data architectures

| Predictive analytics tools

| Data mining tools

| Visualization tools

| Predictive Analytics

| Ethics

| Predictive Customer Insight

Recommended literature for preparation and follow-up:

| Davis K: Ethics of Big Data: Balancing Risk and Innovation

| Grossmann W & Rinderle-Ma S: Fundamentals of Business Intelligence (Data-Centric Systems and Applications), Auflage 2015, Springer, 2015.

| Gupta D: Applied Analytics through Case Studies Using SAS & R, 1st edition, Apress, 2018.

| Han J *et al.*: Data Mining: Concepts and Techniques, Elsevier/Morgan Kaufmann, Amsterdam, 2006.

| Kantardzic M: Data Mining, Wiley, 2011.

| Kelleher JD et al.: Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, MIT Press, 2015

| Kerzner: Project Management: A Systems Approach to Planning, Scheduling and Controlling

| Koster K: International Project Management, Sage Publications Ltd., 2009.

| Kuhn M & Johnson K: Applied Predictive Modelling, Corr. 2nd printing, 2018.

| Marr B: Big Data in Practice, 1st edition, Wiley, 2016.

| Nussbaumer Knaflic C: Storytelling with Data: A Data Visualization Guide for Business Professionals, 1st edition, Wiley, 2016

| Powell B: Microsoft Power BI Cookbook: Creating Business Intelligence Solutions of Analytical Data Models, Reports and Dashboards, Packt Publishing, 2017.

| Rausch P et al.: Business Intelligence and Performance Management: Theory, Systems and Industrial Applications, Springer Verlag, 2013.

| Sherman R: Business Intelligence Guidebook: From Data Integration to Analytics, Morgan Kaufman, 2014.

| Verhoef P et al.: Creating Value with Big Data Analytics, Routledge, 2016.

| Wickham H: R for Data Science, 1st edition, O'Reilly, 2017.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3038 – Data Engineering

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
1	2x	5	Compulsory module	5	- 125 hours, thereof: - 25 in class, - 87,5 private studies, - 12,5 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
None	In all following modules	- Project work - Presentation	- Individual work - Seminar - Team work - Problem-oriented learning	Prof. Dr. Ajinkya Prabhune

Learning objectives

Abstract:

This course can be seen as an engineering tool within a Big Data project.

The engineering basis of Big Data projects is contributed by the storage and provision of big data amounts. A central objective of this module is to acquaint the students with the technical fundamentals of Big Data used for data storage and retrieval as well as to provide them with an overview of typical Big Data architectures. Within this overview, standard technologies and concepts like e.g. No-SQL databases, In-Memory Computing, Hadoop, Map-Reduce as well as distributed data processing are introduced. Big Data architecture vary in dependency of the Problem to be solved; this differences are to be analyzed using practical examples.

Learning outcomes Specialist Competence:

| The students know common Big Data architectures.
| After completion of this module the students are able to distinguish between common Big Data architectures
| They are capable of planning and constructing a complete Big Data architecture for storage and provision of big data amounts in dependence of the application context.
| They evaluate, sort and select Big Data technologies adequately regarding the initial Big Data problem.

Learning outcomes Method Competence:

| The students increase their competences in problem solving.

Learning outcomes Social Competence:

| The students intensify their ability to work in virtual teams and are also capable to use the knowledge and abilities distributed amongst the team to solve a problem in a target-oriented manner.

Learning outcomes Personal Competence:

| The students improve their competence to detect and close gaps in knowledge independently.

Course content

Course content catering the aforementioned competences:

| Hadoop and Map Reduce

| No-SQL databases (Key Value Stores, Graph Databases, Document Stores, Columnar Databases)

| CAP Theorem, BASE Principle

| In-Memory Computing

| Real-time data streams

| Stream Processing, Batch Processing

| Data Warehousing

| Scalability

| Lambda architecture

| Edge Processing

Recommended literature for preparation and follow-up:

| Bengtfort B & Kim J: Data Analytics with Hadoop: An Introduction for Data Scientists, O'Reilly, 2016

| Capriolo E *et al.*: Programming Hive, 1st edition, O'Reilly, 2012

| George L: HBase: The Definitive Guide, 1st edition, O'Reilly, 2011

| Garillot F & Maas G: Stream processing with Apache Spark: Best Practices for Scaling and Optimizing Apache Spark, 1st edition, O'Reilly, 2018

| Grus J: Data Science from Scratch, O'Reilly, 2015

| John T, Misra P: Data Lake for Enterprises: Lambda Architecture for Building Enterprise Data Systems, Packt Publishing, 2017

| Kimball R: The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd edition, Wiley, 2013

| Murthy A & Vivilapalli V: Apache Hadoop Yarn: Moving Beyond MapReduce and Batch Processing with Apache Hadoop 2, 1st edition, Addison Wesley, 2014

| Redmond E & Wilson JR: Seven Databases in Seven Weeks: A Guide to Modern Databases and the NoSQL Movement, Pragmatic Bookshelf, 2012

| Pulipaka G: Big Data Appliances for In-Memory Computing: A Real-World Research Guide for Corporations to Tame and Wrangle Their Data, 2nd edition, High Performance Computing Institute of Technology, 2015

| Sadalage P: NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence, Adisson Wesley, 2009

| White T: Hadoop: The Definite Guide, O'Reilly, 2015

| Recent research literature from peer-reviewed journals

Constructive Alignment
Beginning with the initial problem to construct big data architecture for a specific application scenario, the students develop the required know-how to plan and construct an adequate architecture. For this purpose they first collect the knowledge and skills distributed within the team and are able to detect and close knowledge as well as competence gaps. This module is also accompanied by the analysis of practical examples and exercises concerning the most common big data technologies. Using the examination form practical journal and presentation the students document their learning improvements continually and finally defend them in their presentations.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3039 – Data Management

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)		
1 & 3	2x	15	Compulsory module	8	- 200 hours, thereof: - 40 in class, - 140 private studies, - 20 coaching		

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
none	- Case studies 1 - Case studies 2 - M. Sc. project	- Project work/ development - Presentation	- Individual work - Team work - Exercises - Lecture	Prof. Dr. Barbara Sprick

Learning objectives

Abstract:

The module Data management addresses data acquisition, management and curation.

Through the digitalization of information, implementation of sensors in items of daily use, the mutual communication of systems (Internet of Things) as well as the utilization of social networks, the amount of produced and collected date increases steadily. However, there is a huge heterogeneity in type (time series, text and image files, audio and video files, spreadsheets, etc.) as well as quantity within the data collected in this manner.

For a useful generation of knowledge out of raw data, the latter needs to be cleansed, enhanced, organized, described, etc. in the manner according to the project objective. The main aim of Data management and curation is to model, cleanse and prepare heterogeneous data originating from different sources – this includes structured as well as semi-structured data sources.

The principles of data management are taught in the first semester as well as in Case Studies 1. In the beginning, the students train the foundations of data management in individual exercise units using raw data provided by the SRH University.

New, more complex aspects of data curation are added during the third semester, which enables the students to model, enhance and densify huge data amounts by using metadata knowledge. This is achieved in such a manner so that the data value is increased significantly, thus leading Big Data being transformed into Smart Data. The SRH University provides raw data for exercise purposes which are to be prepared and modelled by the students for a specifically defined application context.

Learning outcomes Specialist Competence:

| After completion of this Module the students know methods and technologies of the management and curation of huge data amounts.
| The students apply methods and technologies for managing and curation of different kinds of huge data amounts.
| They are able to evaluate the quality (i.e. veracity, validity) and the benefit (value) of data regarding a well-defined scenario.
| They may integrate data from different sources and formats.
| They are able to cleanse, homogenize, aggregate and prepare adequately collected data according to a well-defined application context.
| They know the importance of metadata for the value of data amounts and are able to evaluate the significance of specific metadata in context of a certain scenario.

Learning outcomes Method Competence:

| The students know methods and tools for managing and curating data and are able to use a certain spectrum of them.

Learning outcomes Social Competence:

| The students intensify their ability to work in virtual teams and are also capable to use the knowledge and abilities distributed amongst the team to solve a problem in a target-oriented manner.

Learning outcomes Personal Competence:

| The students improve their competence to detect and close gaps in knowledge independently

Course content

Course content catering the aforementioned competences:

| Data identification, verification, cleansing, transformation and integration
| Big Data Variety
| Streaming data
| Batch processing
| Data quality
| Data source categories
| Internet of Things
| Data formatting, scheme-mapping, cleansing
| Time series
| Metadata
| Data Aging, Data Lifecycle Management
| Data profiling
| ETL processes
| syntactic und semantic data transformation

Recommended literature for preparation and follow-up:

| Barton RD: Talend Open Studio Cookbook, Packt Publishing, 2013
| Blokdyk G: Data transformation: A Clear and Concise Reference, CreateSpace Independent Publishing Platform, 2018
| Blokdyk G: Information Lifecycle Management, la …, 2nd edition, CreateSpace Independent Publishing Platform, 2018
| Chambers B & Zaharu M: Spark: The Definitive Guide: Big Data Processing Made Simple, O'Reilly, 2018
| Cielen D & Meysman A: Introducing Data Science, Manning Verlag, 2016
| Garofalakis M & Gehrke J: Data Stream Management: Processing High-Speed Data Streams (Data-Centric Systems and Applications), Springer Verlag, 2016
| Greengard S: The Internet of Things, MIT Press, 2015.

| | Halevy A et al.: Principles of Data Integration, Elsevier LTD, 2012.
| | Komball R & Caserta J: The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data, Kimball Group, 2004
| | Mitchell MN: Data Management Using Stata: A Practical Handbook, Stata Press, 2010
| | Shumway RH & Stoffer DS: Time Series Analysis and Its Applications – With R Examples, 4th edition, Springer, 2017
| | Verborgh R & De Wilde M: Using OpenRefine, Packt Publishing, 2013
| | Recent research literature from peer-reviewed journals

Constructive Alignment

Beginning with the problem to guarantee the quality of the provided raw data in an application context, the students develop the necessary know-how in the field of data management. Starting with collecting the knowledge distributed amongst the team members, the students recognize and closed knowledge gaps by researching and exercising in their respective groups. This module is accompanied by classes providing an introduction to data management methods. The students proof their gain in competences in a project work as well as a final presentation.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3042 – Analytics I

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
1	2x	5	Compulsory module	6	- 150 hours, thereof: - 30 in class, - 105 private studies, - 15 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
none	Case Studies	- Written exam	- Individual work - Team work - Exercises	Prof. Dr. Herbert Schuster

			- Lecture	

Learning objectives

Abstract:

During this module the students learn the essential tools and methods of inductive and descriptive statistics. They are able to perform first analyses individually using standard tools (SAS, SPSS, etc.). Hereby the students understand the applicability, prerequisites and the interpretation purposes of the diverse statistical methods taught during this module.

Learning outcomes Specialist competence:

| The students are able to use the relevant statistical concepts and methods adequately. They understand methods in context of their function and are capable to use them in a problem-solving manner.
| They may execute more complex analyses and understand to evaluate the gained results analytically.
| They are capable to adapt their gained knowledge and competence to solve new, unknown projects in an objective-oriented manner.

Learning outcomes Method competence

| The students are able to apply their theoretical competences practically on selected software systems.
| The may evaluate the adequate usage of these competences.

Learning outcomes Social and personal competence:

| The students are capable to analyze and classify problems individually as well as in teams and may develop user-centered solutions based on the analysis and classification results.

Course content

Course content catering the aforementioned competences:

Descriptive statistics

| Introduction of key concepts of logic and statistics
| Principles of statistics
| Classification of variables
| Measure of central tendency
| Dispersion parameters
| Simple and multiple regression and correlation analysis
| Modelling concepts
| Cluster processes

Inductive statistics

| Introduction of combinatorics
| Principles of the theory of probabilities
| Theoretical distributions
| Theory of sampling and estimation methods

All chapter are taught in a methodologically-theoretically as well in application-oriented manner. Therefore problems, on which the introduced methods specifically could be applied, as well as a complete application are used.

Recommended literature for preparation and follow-up:

| Bruce, P & Bruce A: Practical Statistics for Data Scientists: 50 Essential Concepts, 1st edition, O'Reilly, 2017.
| Chang W: R Graphics Cookbook, 1st edition, O'Reilly, 2012
| Field et al.: Discovering Statistics Using R, Sage Publications, 2012
| Grolemund G: Hands-On Programming with R, 1st edition, O'Reilly, 2014
| James G et al.: An Introduction to Statistical Learning: with Applications in R, Corr. 7th printing, 2017
| Izenman AJ: Modern Multivariate Statistical Techniques, Corr. 2nd Printing, Springer, 2013
| Schumacker RE: Using R With Multivariate Statistics, Sage Publications, 2015
| Sheather SJ: A Modern Approach to Regression with R, Springer, 2009
| Shumway RH & Stoffer DS: Time Series Analysis and Its Applications – With R Examples, 4th edition, Springer, 2017
| Wickham H: ggplot2: Elegant Graphics for Data Analysis, 2nd edition, Springer, 2016
| Wickham H & Grolemund G: R for Data Science, 1st edition, O'Reilly, 2017

Constructive Alignment
The students are examined using a written examination. Through this examination form the evaluation of the basic vocabulary as well as the application of the theoretical principles on well-defined scenarios. Another objective of this examination is the transfer of the principles learned on new, complex problems.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3043 – Analytics II

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; for exceptions view appendix 2b of the examination regulations)
2	2x	5	Compulsory module	6	- 150 hours, thereof: - 30 in class - 105 private studies, - 15 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
- Analytics I	- In the case studies - Master thesis project	- Oral examination - Test	- Problem-oriented work - Seminar - Exercises	Prof. Dr. Herbert Schuster

Learning objectives

Abstract:

The students know the challenges posed by Big data on the structured data acquisition and their processing into information being helpful for making business-relevant decisions. They are capable of structuring complex problems and performing systematic research work. Based on huge data amounts they may analyse and prepare information to generate user-centered knowledge. They are able to select the adequate data mining techniques to solve specific business-relevant problems and to visualize the gained results appropriately. Finally the results are evaluated critically regarding their validity.

Learning outcomes Specialist competence:

| The students know the essential methods for the procedural steps of the preparation of data mining methods: preparation of raw data, structuring and refining.
| They are capable to perform and apply the most important methods of context analyses
| They are able to perform more complex analyses and to evaluate the results in a functional way.
| They may critically reflect the validity of the results regarding qualitative as well as quantitative aspects.
| They know the most important web mining tools and are capable to apply them in relevant practical exercises.
| They identify state of the art concept to visualize data mining results.

Learning outcomes Method competence:

| The students may apply correctly the introduced methods by solving specific problems and interpreting the results adequately.
| They know the most essential methods in data mining analysis.
| They are able to apply the methods learned by using specific software solutions and may critically reflect the result's validity.

Learning outcomes Social and personal competence

| The students are capable to analyze the methods used as well as the results gained in their entity and evaluate them benefit-oriented during a business-specific decision-making process.

Course content

Course content catering the aforementioned competences:

| Introduction of data mining methods

| Data mining as a process

| Methods 1: Decision trees

| Methods 2: Association rules and sequence patterns

| Methods 3: Classification theories

| Methods 4: Variance analysis, factor analysis, discriminant analysis

| Methods 5: time series analysis

All methods are applied on defined examples via the use of standard analysis systems.

Recommended literature for preparation and follow-up:

| Buttcher S *et al.*: Information Retrieval: Implementing and Evaluating Search Engines, Reprint, MIT Press, 2016

| Gorsuch RL: Factor Analysis: Classic Edition (Routledge Classic Editions), 2., new edition, 2014

| Han J *et al.*: Data Mining: Concepts and Techniques, Elsevier/Morgan Kaufmann, Amsterdam, 2006

| Huberty CJ & Olejnik S: Applied MANOVA And Discriminant Analysis, 2nd edition, Wiley-Interscience, 2006

| Kantardzic M: Data Mining, Wiley, 2011

| Liu B: Web Data Mining: Exploring Hyperlinks, Contents and Usage Data (Data-Centric Systems and Applications, 2. Auflage, Springer, 2011

| Marsland S: Machine Learning - An Algorithmic Perspective, CRC Press, 2009

| Matloff N: Statistical Regression and Classification: from Linear Models to Machine Learning (Texts in Statistical Science), Chapman & Hall, 2017

| Mitchell TM: Machine Learning, McGraw-Hill, 1997

| Provost F & Fawcett T: Data Science for Business: What you need to know about data mining and data-analytic thinking, 1st edition, O'Reilly, 2013

| Runkler TA: Data Analytics – Models and Algorithms for Intelligent Data Analysis, 2nd edition, Springer Vieweg, 2016

| Sahai H & Ageel MI: The Analysis of Variance: Fixed, Random and Mixed Models, reprint, Birkhäuser, 2012

| Singaravelu VVM: Sequence Association Rule With Pattern Constraints In Web Usage Mining: Data Stream Mining, LAP LAMBERT Academic Publishing, 2011

| Sullivan W: Decision Tree and Random Forest; Machine Learning And Algorithms: The Future Is Here!, CreateSpace Independent Publishing Platform, 2018

| Witten IH & Eibe F: Data Mining: Practical Machine Learning Tools and Techniques, 4th edition, Morgan Kaufman, 2016

The students proof their application-oriented knowledge and competences by solving well-defined problems and exercises during an oral examination. This form of examination also evaluates the student's interpretation capabilities.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3044 – Analytics III

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
3	2x	10	Compulsory module	6	- 150 hours, thereof: - 30 in class, - 105 private studies, - 15 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
- Analytics I	- In the case studies - M. Sc. project	- Oral examination - Test	- Problem-oriented learning - Seminar - Exercises	Prof. Dr. Herbert Schuster

Learning objectives

Abstract:

During this module the students learn the key methods of predictive analytics and are capable of using the common standard tools (e.g. SAS, SPSS, BO, IBM Operation Analytics) to perform analyses independently and critically evaluate the results in terms of validity and adequate use. They are able to apply their gained competences in new, complex situations in an objective-oriented manner.

Learning outcomes Specialist competence:

| The students are able to use the functional terms learned during this module adequately as well as applying the methods in a practice-oriented way.

| They are able to perform more complex analyses and evaluate the results in a functional way.

| They may critically reflect the validity of the results regarding qualitative as well as quantitative aspects.

Learning outcomes Method competence:

| They are able to apply the methods learned by using specific software solutions and may critically reflect the result's validity.

Learning outcomes Social and personal competence:

| The students are capable to analyze and classify problems individually as well as in teams and may develop user-centered solutions based on the analysis and classification results.

Course content

Course content catering the aforementioned competences:

Analytics III is a practice-oriented module, which is based on the SAS course „Advanced Analysis in a Big Data World, HAW Predictive Analytics and Social Media Analytics". The module is held in cooperation with SAS.

| Neural networks
| Support Vector Machines
| Opening of Black Box for neural networks and SVMs
| Regression trees
| Ensemble methods
| Rule types and alternative rule descriptions
| Bayes network classifier
| Survival time analysis
| Social Networks: Learning and interference
| Surveillance and back testing of analytical methods

Recommended literature for preparation and follow-up:

| Crowder MJ: Multivariate Survival Analysis and Competing Risks, Routledge, 2016

| Flach P: Machine Learning: The Art and Science of Algorithms that make Sense of Data, New edition, Cambridge University Press, 2012

| Han J *et al.*: Data Mining: Concepts and Techniques, Elsevier/Morgan Kaufmann, Amsterdam, 2006

| Kantardzic M: Data Mining, Wiley, 2011

| Khan GF: Creating Value With Social Media Analytics: Managing, Aligning, and Mining Social Media Text, Networks, Actions, Location, Apps, Hyperlinks, Multimedia, & Search Engines Data, 1st edition,

CreateSpace Independent Publishing Platform, 2018.

| Kuhn M: Applied Predictive Modeling, Corr. 2nd Printing, Springer, 2018

| Kelleher JD et al.: Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, MIT Press, 2015

| Liu B: Web Data Mining: Exploring Hyperlinks, Contents and Usage Data (Data-Centric Systems and Applications, 2. Auflage, Springer, 2011

| Ma X: Using Classification and Regression Trees: A Practical Primer, Information Age Publishing, 2018

| Marsland S: Machine Learning - An Algorithmic Perspective, CRC Press, 2009

| Murty MN & Raghava R: Support Vector Machines and Perceptrons: Learning, Optimization, Classification, and Application to Social Networks, 1st edition, 2016.

| Rashid T: Make your own Neural Network, 1st edition, CreateSpace Independent Publishing Platform, 2016

| Runkler TA: Data Analytics – Models and Algorithms for Intelligent Data Analysis, 2nd edition, Springer Vieweg, 2016

| Schwarz T: Big Data im Marketing: Chancen und Möglichkeiten für eine effektive Kundenansprache, Haufe-Lexware, 2015

| Witten IH *et al.*: Data Mining : Practical Machine Learning, Tools and Techniques, 3rd edition, Elsevier, 2011

Constructive Alignment

The students proof their application-oriented knowledge and competences by solving well-defined problems and exercises during an oral examination. This form of examination also evaluates the student's interpretation capabilities.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3045 – Data Storytelling and Communication

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; for exceptions view appendix 2b of the examination regulations)
1,2,3	2x	20	Compulsory module	12	- 300 hours, thereof: - 60 in class, - 210 private studies, - 30 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
none	- In case studies - M. Sc. project	- Portfolio exam - Presentation	- Seminar - Lecture - Exercises - Team work	Prof. Dr. Ajinkya Prabhune

Learning objectives

Abstract:

The module "Data Storytelling and Communication" considers effective communication of insights regarding the original problem. These insights are often described using dashboards, infographics, etc. The skills required for this are introduced during the courses held during the first and second semester. Information has to be interpreted, set in context with the original problem and analyzed data as well as to be deduced into insights, so that businesses may benefit from these visualizations and develop further actions. It is necessary to tell a "Data story" based on the original problem, reaching from data choice and analysis to insights and insight-derived actions. A central aspect of data storytelling is to communicate insights appropriate to the target group.

This module introduces the principles of traditional storytelling as well as those of data-driven storytelling.

Learning outcomes Specialist competence:

| The students are able to use data visualizations for interactive storytelling enabling and supporting the exploration of analysis results as well as the derivation of new problems.
| They may visualize and communicate analysis results in a target group-oriented way.

Learning outcomes Method competence:

| After completion of this module the students know tools supporting interactive storytelling (e.g. GeoTime Stories, Tableau Public) and are able to use them in a target-oriented manner.
| They are capable to prepare insights according to their target group and decisions regarding the original problem

Learning outcomes Social competence:

| They improve their communication abilities
| They are capable of recognizing the needs of a target group so that they prepare information and communicate insights properly.

Learning outcomes Personal competence:

-

Course content

Course content catering the aforementioned competences:

| Theories of story setup, structures and intentions
| Explorative data analysis
| Effectively written communication
| Effectively visual presentations
| Design principles (Forms, colors, etc.)
| Human perception
| Story development and compaction
| Interplay between narrative and visual communication
| interactive Storytelling
| Infographics, dashboards, etc.
| Tools: Tableau, Illustrator, R/ggplot, Photoshop, GeoTime Stories, Tableau Public

Recommended literature for preparation and follow-up:

| Alexander B: The New Digital Storytelling: Creating Narratives with New Media, ABC-Clio, 2011
| Berinato S: Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations, Harvard Business Review Press, 2016
| DeBarros A: Practical SQL: A Beginner's Guide to Storytelling with Data, No Starch Press, 2018
| Evergreen SDH: Effective Data Visualization: The Right Chart for the Right Data, Sage Pubn, 2016
| Foreman JW: Data Smart: Using Data Science to Transform Information into Insight, 1st edition, Wiley, 2013
| Marr B: Big Data: Using SMART Big Data, Analytics and Metrics to Make Better Decisions and Improve Performance, 1st edition, Wiley, 2015
| Nussbaumer Knaflic C: Storytelling with Data: A Data Visualization Guide for Business Professionals, 1st edition, Wiley, 2016
| Provost F & Fawcett T: Data Science for Business: What you need to know about data mining and data-analytic thinking, 1st edition, O'Reilly, 2013
| Wong DM: The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures, Reprint-Auflage, Ww Norton & Co, 2014

Constructive Alignment
The students analyze different story examples and subsequently develop own stories based on role-plays. The methodological spectrum thereby reaches from data selection and visualization to interpretations for different target groups. This work is performed individually as well as in teams. Finally the students or their respective groups present their developed stories to each other and subsequently critically reflect the results. The combined examination form consisting of presentation and portfolio suit the module intention adequately because there is a continuous documentation and evaluation of the students' improvements in competences as well as of the status of the data story..

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module 3049 – Privacy, Ethics and International Law

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Student's workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)		
2,4	2x	10	Compulsory module	7	- 175 hours, thereof: - 35 in class, - 122,5 private studies - 17,5 coaching		

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
none		- Written exam - Exposé	- Problem-oriented learning - Seminar - Colloquium	Prof. Dr. Anke Schuster

Learning objectives

<u>Abstract:</u>

During this module the students develop fundamental knowledge about privacy, ethics and the judicial aspects

in context of data analysis. They generate an awareness of ethically relevant problems and they are able to evaluate individual, social and institutional actions in socio-technical situations (e.g. based on privacy law). Additionally, they learn to impose privacy requirements through organizational-technical measures.

This module comprises two courses: "Privacy and its realization" during the second and "Ethics and Law" during the fourth semester.

Learning outcomes Specialist competence:

| The students are able to examine contexts of origin and effects from an ethical perspective and may apply ethical and privacy concepts on defined examples of socio-technical scenarios.
| They know the prerequisites of a transparent, informed approval as well as the prerequisites of data transfer and may derive consequences for big data projects.
| They are capable of reproducing and applying the principles of data curation and utilization according to national and international law
| They know and exert the relevant privacy laws, regulations and strategies.

Learning outcomes Method competence:

| The students know and target-orientedly apply organizational as well as technical measures to impose privacy and personal rights

Learning outcomes Social and personal competence:

| The students may analyze and evaluate well-defined problems independently.
| They a are able discuss in a functional and scientific way.

Course content

Course content catering the aforementioned competences:

Ethics and international law

| Terminology of ethics, business ethics
| Ethics within the technical civilization/occupations
| Individual and institutional ethics
| Ethical codices for computer scientists
| Ethics within an interconnected world
| Lawful actions and conflict of interests
| Rights of the persons affected

- International data processing and jurisdiction
- Principles of appropriation and approval requirements
- Regimentation in big data inquirys
- Contracts regarding data and data analyses
- German privacy, internet and communication laws (Bundesdatenschutzgesetz, Telemediengesetz, Telekommunikationsgesetz)
- Data transfer within a business and places outside the EU

Privacy and its enforcement

- Principles of privacy law
- Data separation
- Technologies to enforce privacy requirements
- Organizational measures
- Anonymization and pseudonymization
- Application scenarios
- Risks caused by data aggregation
- Misuse of data

Recommended literature for preparation and follow-up:

- Anderson M: Machine Ethics, reprint, Cambridge University Press, 2018
- Craig T & Ludloff ME: Privacy & Big Data, 1st edition, O'Reilly, 2011
- Davis K: Ethics of Big Data: Balancing Risk and Innovation, 1st edition, O'Reilly, 2012
- Floridi L: The ethics of Information, reprint, Oxford University Press, 2015
- Holmes RL: Introduction to Applied Ethics, Bloomsbury Publishing, 2018
- Kuhlen R: Informationsethik, Konstanz, 2004.
- Kuner C: Transborder Data Flow Regulation and Data Privacy Law, Oxford University Press, 2013
- Lane J: Privacy, Big Data, and the Public Good: Frameworks For Engagement, Cambridge University Press, 2014
- Stamatellos, G: Computer Ethics. A global perpective, Sudbury, 2007.
- Zimmer M & Kinder-Kurlanda K: Internet Research Ethics for the Social Age: New Challenges, Cases, and Contexts, Peter Lang Publishing, 2017

Constructive Alignment

During the course "Ethics and law" the students learn, analyze and discuss ethical and judicial aspects in context of big data and data analysis through well-defined practical examples as well as presentations. The gained theoretical competences in the actual privacy laws and regimentations are evaluated through a written examination. This form of examination additionally enables the students to reproduce, apply and discuss judicial aspects of privacy law on well-defined examples and scenarios.

The course "Privacy and its enforcement" enables the students to develop technical and organizational measures to enforce privacy and personality laws in big data projects and data analyses. The evaluation of the student's competence is performed via case work.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module A-1003 – Master Thesis Project

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Workload (1 ECTS appropriates 25 h; for exceptions view appendix 2b of the examination regulations)		
4	2x	20	Compulsory module	18	- 450 hours, thereof: - 0 in class, - 405 private studies, - 45 coaching		

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
Scientific work and ethics		- Project work	- Problem-oriented learning - Data research, analysis and processing	Prof. Dr. Barbara Sprick

Learning objectives

Abstract:

The students understand the fundamental aspects of scientific work and are able to structure and exert individually the cognitive process from the original problem to the systematic answering of a well-defined scientific problem. They know the essential methods and tools for the development of a scientific work and

may critically reflect the results. They are capable of a qualitative as well as quantitative evaluation of method use. The students may work independently on a scientific big data and business analytics problem using common scientific methods and gain new insights.

Learning outcomes Specialist and method competence:

| The students are able to structure their scientific cognitive process. They yield results according to the criteria of good scientific work (i.e. objectivity, validity and reliability)
| They are capable of giving proper qualitative as well as quantitative judgements regarding the adequate use of scientific methods
| They may critically evaluate and reflect the gained results and method.
| They intensify functional and scientific-methodological competences learned during the master program over the defined problem of the master thesis project.
| They are able to transfer the knowledge of "Scientific work and ethics" on the master thesis project.
| They are competent to lead and moderate a functional scientific discussion to analytically-critically reflect scientific results and use of methods.

Learning outcomes Social competence:

| The students are able to evaluate results, gain in insights on a functional basis and may verbalize constructive feedback.
| They are capable of leading a functional discussion to gain insights.

Learning outcomes Personal competence:

| The students are able to perform research work systematically and independently as well as to reflect insights using iterative thinking processes.
| They are competent to structure the scientific cognitive process of the master thesis project regarding scheduling, systematic structuring and gaining of insights.

Course content

Course content catering the aforementioned competences:

Scientific work and writing

| Scientific quality criteria
| Scientific methods
| Criteria to evaluate scientific works
| Research, classification and evaluation of scientific literature
| Scientific writing
| Leading and moderation of scientific discussions

Recommended literature for preparation and follow-up:

| Heard SB: The Scientist's Guide to Writing: How to Write More Easily and Effectively throughout Your Scientific Career, Princeton University Press, 2016
| Skern T: Writing Scientific English: A workbook, 2nd edition, UTB, 2011

Links

| Google Scholar
| DBLP
| IEEE Computer Society
| IEEE TVCG camera ready document guidelines

Literature management

| Citavi

Constructive Alignment

Module is completed by the master thesis project work.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module xxxx – Big Data Programming I

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
1	2x	20	Compulsory module	6	- 150 hours, thereof: - 30 in class, - 105 private studies, - 15 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
-	- All further modules	- Project work - Presentation	- Lecture - Learning Lab - Theoretical and practical exercises	Prof. Dr. Barbara Sprick

Learning objectives

Abstract:

The main intention of this module is to provide the students with programming languages being fundamental for machine-learning and Big Data Analytics processes, e.g. Python and R. Additionally the students get insight into the adequate handling of databases like MongoDB and NoSQL databases. Furthermore, the students are introduced into basic machine-learning using TensorFlow sessions as well as neural networks. Didactically, the module is composed of a mix of lectures, practical lab sessions and exercises to ensure the student's methodological progress.

Learning outcomes Specialist and Method competence:

| The students know the fundamentals of the Python and R programming languages and are able to develop and implement more complex programs independently.
| They are able to adequately handle MongoDB and NoSQL databases.
| For a given instance they are competent to compare the concepts and argue for an adequate choice in the field of machine-learning using TensorFlow sessions as well as neural networks.
| They are able to develop, implement and apply python programming for TensorFlow sessions.
| They are able to independently develop simple machine-learning procedures

Learning outcomes Social and personal competence:

| The students improve their ability to analyse problems, to break large problems down into digestible portions.

| By presenting their own and other's work, they also improve their communication skills.

Course content

Course content catering the aforementioned competences:

| Introduction to the Python and R programming language

| Source code management, revision and branch and version management, refactoring

| Software documentation and tools

| Test driven development and architecture

| Fundamentals of pipeline programming

| Agile development strategies of intelligent systems

| Client Libraries

    | MongoDB

    | SQL-Datenbank

| Fundamentals of TensorFlows

    | General overview, functionality and handling

    | TensorFlows in popular applications (e.g. Google)

    | Strengths and drawbacks

    | TensorFlow sessions

    | TensorFlow in machine-learning processes

    | Visualizations using TensorFlow (Graph Management, Node Value, Linear Regression)

    | TensorFlow across devices and servers (operations on single and multiple devices on sessions using Python)

| Neural networks

    | General intention to create neural networks

    | Saving and restoring models

    | Artificial neural networks

    | Neural networks in machine-learning-processes

    | Visualization and Training curves

    | Application of simple machine-learning algorithms on in programming I

Recommended literature for preparation and follow-up:

| Géron A: Hands-On Machine Learning with Scikit-Learn & TensorFlow – Concept, Tools, and

Techniques to Build Intelligent Systems, 1st edition, O'Reilly, 2017

| Goodfellow I *et al.*: Deep Learning (Adaptive Computation and Machine Learning), MIT Press, 2016

| Gulli A & Kapoor A: TensorFlow 1.x Deep Learning Cookbook: Over 90 Unique Recipes To Solve Artificial-Intelligence Driven Problems With Python, Packt Publishing, 2017

| Martin RC: The Clean Coder: A Code of Conduct for Professional Programmers, 1st edition, Prentice Hall, 2011

| Martin RC: Clean Architecture: A Craftsman's Guide to Software Structure and Design, 1st edition, Prentice Hall, 2017

| Pattanayak S: Pro Deep Learning with TensorFlow: A Mathematical Approach to Advanced Artificial Intelligence in Python, 1st edition, APress, 2017

| Ramalho, L: Fluent Python, 1st edition, O'Reilly, 2015

| Wickham H & Grolemund G, R for Data Science, 1st edition, O'Reilly, 2017

Constructive Alignment

Participants will practise project management methods in a realistic project setting and therefore also enhance their social- and self-competencies by working in teams. They enhance their software development skills by planning and realizing a project using modern software engineering concepts and tools. The examination form project work was chosen to evaluate the student's methodological abilities in programming and machine-learning processes. Additionally the project work may be used to proof the student's advanced scientific abilities. During the presentation, the students are given the opportunity to present and defend their results gained during the project work in front of their class-mates.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module xxxx – Big Data Programming II

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
3	2x	20	Compulsory module	8	- 200 hours, thereof: - 40 in class, - 140 private studies, - 20 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
- Big Data Programming I	- All further modules	- Project work - Presentation	- Lecture - Learning Lab	Prof. Dr. Barbara Sprick

			- Theoretical and practical exercises	

Learning objectives

Abstract:

The main intention of this module is to enable the students to gain fundamental as well as advanced knowledge in the field of artificial intelligence. Additionally, a profound knowledge base in Natural Language Processing will give students the opportunity to apply and evaluate NLP-dependent processes like e.g. Twitter. Finally the students abstract their theoretical and methodological competences in the fields of AI and NLP to independently concept, develop and implement a messaging system on their own.

Learning outcomes Specialist and Method competence:

| The students will understand important principles and application of AI
| They are able to apply these principles on well-defined examples.
| The students are capable of the application of the most important NLP processes.
| The students are able to transfer NLP principles to the NLPs for Communication
| They are able to analyze and critically evaluate NLP application (e.g. Twitter messages)
| They may use their specialist and methodological competences in the field of NLP as well as their scientific competences to independently develop a complex messaging system

Learning outcomes Social and personal competence:

| The students improve their ability to analyse problems, to break large problems down into digestible portions.
| By presenting their own and other's work, they also improve their communication skills.

Course content

Course content catering the aforementioned competences:

| Principles of Artificial Intelligence

| Natural Language Processing

   | Language Models

   | Text Classification

   | Information Retrieval

   | Information extraction

   | Summary, Bibliographical and historical notes

| Natural language for Communication

   | Phrase Structure Grammars

   | Syntactic Analysis

   | Augmented Grammars and Semantic Interpretation

   | Machine Translation

   | Speech recognition

| Natural Language Processing in a pracitical context

   | Analysis of Twitter messages

   | Development of a messenger system

Recommended literature for preparation and follow-up:

| Bengfort B: Applied Text Analysis with Python: Enabling Language Aware Data Products with Machine Learning, O'Reilly, 2018.

| Millstein F: Natural Language Processing With Python: Natural Language Processing Using NLTK, CreateSpace Independent Publishing Platform, 2018.

| Porsiel J: Machine Translation: What Language Professionals Need to Know, BDÜ Fachverlag, 2017.

| Stent A & Bangalore S. Natural Language Generation in Interactive Systems, Cambridge University Press, 2014.

| Stephens-Davidowitz S: Everybody lies: What the Internet Can Tell Us About Who We Really Are, Bloomsbury, 2018

| Yu D & Deng L: Automatic Speech Recognition: A Deep Learning Approach (Signals & Communication Technology)

| For Python manuals and tutorials see www.python.org

| For Scikit learning manuals and tutorials see www.scikit-learn.org

| For TensorFlow manuals and tutorials see www.tensorflow.org

Constructive Alignment

The examination concluding this module will be a combination of a theoretical and practical assessment. In the theoretical assessment, the students will show their knowledge and understanding of important threats and weaknesses of the common concepts, methods and techniques of AI as well as NLP. In the practical assessment, the students will show their ability to analyse the an NLP process on a well-defined example, decide on and argue for appropriate solutions for a particular NLP problem and independently a complex NLP-dependent process on their own using their specialist, methodological and scientific competences.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module xxxx – Information Systems

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
2	2x	20	Compulsory module	8	- 200 hours, thereof: - 40 in class, - 140 private studies, - 20 coaching

Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator
	- All case studies	- Project work	- Lecture - Tutorials - Seminar - Project	Prof. Dr. Barbara Sprick

Learning objectives

Abstract:

This module aims at qualifying the students to accomplish complex projects in the area of modern information systems for structured, semi-structured and unstructured data, with various types of data-models: e.g. relational databases, graph-databases, key-value stores, columnar databases, or document databases.

In small groups of 2-4 students plan, design and realise a comprehensive database project using adequate instances of modern NoSQL information systems.

Learning outcomes Specialist and Method competence:

| The students know various instances of modern information systems for structured, semi-structured and unstructured data with their respective data models, e.g. relational, graph-based, columnar, key-value, or document data models.

| They understand the differences, strength and weaknesses of various data models and corresponding data bases and can use them in an adequate way for a particular use case.

| They can describe and classify the various concepts. For a given problem instance, they can compare the concepts and argue for an adequate choice.

| The students understand the theoretical concepts behind NoSQL databases.

Learning outcomes Social and personal competence:

| The students improve their ability to analyze problems, to break large problems down into digestible portions.

| By presenting their own and other's work, they also improve their communication skills.

Course content

Course content catering the aforementioned competences:

| Distributed information systems
  | Classes and architectures of distributed information systems
  | Scheme design
  | Distributed query evaluation
  | CAP / BASE
| NoSQL Databases
  | Definition, Classification/Categorization
  | Basic concepts (Data- and Storage models, Query models, query languages, index structures)
  | Basic technologies (scalability, distribution, parallel query processing, MapReduce, Transaction handling, replication)
| Information systems for semi-structured and unstructured data
| Different data models: graph-based, document stores, key-value stores, columnar DBs and relational DBs

Recommended literature for preparation and follow-up:

| Harrison G: Next Generation Databases: NoSQL, NewSQL and Big Data, 1st edition, Apress, 2015.

| Lake P & Drake R: Information Systems Management in the Big Data Era (Advanced Information and Knowledge Processing), 1st edition reprint, 2016.

| Ploetz A & Kandhare D: Seven NoSQL Databases in a Week, Packt Publishing, 2018.

| Tanenbaum AS & van Steen M: Distributed Systems: Principles and Paradigms, 2nd edition, CreateSpace Independent Publishing Platform, 2016.

Constructive Alignment

This module aims at qualifying the students to accomplish complex projects in the area of modern information systems for structured, semi-structured and unstructured data. Basic knowledge about modern information systems will be presented in a lecture. The students will deepen their understanding by planning and realizing their own project.

In the project report and presentation of their project, the students will prove their ability to pursue a project with a modern information system. In the seminar they will demonstrate their understanding of the area as well as their ability to compare concepts and argue for adequate choice of a modern information system.

SRH University Heidelberg, M. Sc. full-time program Big Data & Business Analytics

Module xxxx – Analytics IV: Machine Learning

Semester	Annually offered	Duration (weeks)	Type	ECTS points	Workload (1 ECTS appropriates 25 h; For exceptions view appendix 2b of the examination regulations)
4	2x	20	Compulsory module	8	- 200 hours, thereof: - 40 in class, - 140 private studies, - 20 coaching
Requirements for enrolment	Applicability	Type(s) of exam	Teaching and learning methods	Module coordinator	
- Big Data Programming I - Analytics I-III	- All following modules	- Examination (50%), - Project work + presentation (50%)	- Lecture - Tutorials	Prof. Dr. Ajinkya Prabhune	
Learning objectives					
Abstract:					

During this module, the students deepen their specialist and methodological abilities in machine learning processes gained in Big Data Programming I. The knowledge of neural networks is deepened and the methodological focus is set on the evaluation and choice of adequate machine learning methods for a given machine-learning problem.

Learning outcomes Specialist and Method competence:

| Students will improve their competences in neural networks and are able to choose the adequate machine-learning method depending on the type of neural network
| Learning outcomes Social and personal competence:
| Beside these basic application-building skills, students will be able to understand and manage team-based projects in application development.
| They train necessary leadership competencies alongside the second project-centered part.
| At the end of the course, students are capable to program as well as improve their practical teamwork-experience.
| With the final presentation students also show their qualification in presentation of scientific, methodical and technological outcomes.
| The students improve their ability to analyze problems, to break large problems down into digestible portions.
| By presenting their own and other's work, they also improve their communication skills.

Course content

Course content catering the aforementioned competences:

| Machine learning
  | Repetition of the fundamentals teached in Big Data Programming I
  | Application of machine learning processes in neural networks
    o Multi-Layer Perceptrons
    o Deep neural networks
  | Modulation of neural networks
    o Fine tuning of neural network parameters
    o Modulation of neurons per layer
    o Activation functions
  | Evaluation and choice of adequate machine-learning methods
    o Models from other frameworks
    o Tweaking, dropping, replacement of upper neural layers
    o Freezing of lower neural layers

- o Caching of frozen lower neural layers
- o Pre-training (fixed and trainable weights)

Recommended literature for preparation and follow-up:

| Géron A: Hands-On Machine Learning with Scikit-Learn & TensorFlow – Concept, Tools, and Techniques to Build Intelligent Systems, 1st edition, O'Reilly, 2017.

| Allaire JJ: Deep Learning with R, Manning Publications, 2018

| Alpaydin E: Introduction to Machine Learning, 2nd edition, MIT press, 2009.

| Alpaydin E: Machine Learning: The New AI, MIT Press, 2016

| Bishop CM: Pattern Recognition and Machine Learning, corr. 2nd printing, 2011

| Chollet F: Deep Learning with Python, Manning Publications, 2017

| Christian B & Griffiths T:Algorithms To Live By: The Computer Science of Human decisions, Harper Collins, 2017

| Flach P: Machine Learning: The Art and Science of Algorithms that make Sense of Data, New edition, Cambridge University Press, 2012

| Géron A: Hands-On Machine Learning with Scikit-Learn & TensorFlow – Concept, Tools, and Techniques to Build Intelligent Systems, 1st edition, O'Reilly, 2017

| Goodfellow I *et al.*: Deep Learning (Adaptive Computation and Machine Learning), The MIT Press, 2016

| Kelleher JD et al.: Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies, MIT Press, 2015

| Kubat M: Introduction to Machine Learning, 1st edition, Springer, 2011

| Marsland S: Machine Learning: An Algorithmic Perspective, 2nd new edition, 2014

| Murphy KP: Machine Learning: a Probabilistic Perspective, MIT Press, 2012

| Pattanayak S: Pro Deep Learning with TensorFlow: A Mathematical Approach to Advanced Artificial Intelligence in Python, 1st edition, APress, 2017

| Patterson J & Gibson A: Deep Learning: A Practitioner's Approach, 1st edition, O'Reilly, 2017

| Rashid T: Make your own Neural Network, 1st edition, CreateSpace Independent Publishing Platform, 2016

| Russell S, Norvig P: Artificial Intelligence: A modern Approach, Global Edition, 3rd edition, Addison Wesley, 2016.

Constructive Alignment

Participants will study and learn the basic approaches of machine learning in theory as well as in a practical respect. The theoretical part will be structured along special aspects of machine learning and neural networks. In practical sessions, this knowledge has to be used in order to get experience in conducting and bringing to conclusion a typical application-centred project. Social skills and self-management competencies will be enhanced by working in teams under real-world conditions. The result of the individual work (paper) as well as project team work will be presented and evaluated.