Module Handbook

Data and Discourse Studies – Computer Science

Technische Universität Darmstadt Department of Computer Science

Module Handbook Data and Discourse Studies – Computer Science 1



Module Handbook Data and Discourse Studies – Computer Science

Technische Universität Darmstadt

Fachbereich Informatik

Hochschulstr. 10

64289 Darmstadt

Redaktion

Dipl.-Inform. Tim Neubacher

Stand: 13.09.2018

| Modu | le desc | ription | | | | | | | |
|------------|-----------------------|--------------------|-------------------------|--------------|---------------|-----------|------------|--------------|--------------|
| Modu | le | | | | | | | | |
| Inform | nation I | Management | | | | | | | |
| Modu | le | Credit points | Hours required | Sel | f-study | Mod | ule | Recurren | ce frequency |
| No. | | 5 CP | 150 h | 105 | 5 h | dura | tion | Summer s | semester |
| 20-00-0015 | | | | | | 1 Ser | nester | | |
| Langu | Language | | Per | son responsi | ble fo | r the mod | ule | | |
| Germa | nan | | | Vic | e Chairpersor | n, Acad | demic Affa | airs | |
| 1 | Courses in the module | | le | | | | | | |
| | Cours | se Course title | 9 | | Workload (| CP) | Type of | lecture | Contact |
| | No. | | | | | | | | hours per |
| | | | | | _ | | | | week |
| | 20-00 |)- Information | n Management | | 5 | | Integrat | ed course | 3 |
| | 0015 | -IV | | | | | | | |
| 2 | Cont | ent | | | | | | | |
| - | Basic | concepts of info | ormation managem | nent: | | | | | |
| | | | | | | | | | |
| | Conc | epts of informat | ion systems | | | | | | |
| | Infor | mation storage/ | retrieval, searching | g, bro | wsing, declar | ative a | iccess and | l access via | explicit |
| | navig | ation | | | - | | | | |
| | | | | | | | | | |
| | Quali | ty characteristic | s: | | | | | | |
| | Cons | istency, scalabili | ty, availability, relia | ability | , | | | | |
| | Data | | | | | | | | |
| | Data | Modeling: | | | | | | | |
| | Conc | eptual data mod | lels (ER / UML stru | cture | diagrams) | | | | |
| | Conc | eptual design | | | U | | | | |
| | Oper | ational Models (| relational model) | | | | | | |
| | Марр | ping the concept | ual to the operation | onal n | nodel | | | | |
| | | | | | | | | | |
| | Relat | ional model: | | | | | | | |
| | Oner | ators | | | | | | | |
| | Relat | ional algebra | | | | | | | |
| | Relat | ional calculi | | | | | | | |
| | Impa | ct on query lang | uages based on rel | ation | al algebra an | d relat | ional calc | ulus | |
| | Desig | n theory and no | rmalization | | - | | | | |
| | | | | | | | | | |
| | Quer | y languages: | | | | | | | |
| | 501 (| in detail) | | | | | | | |
| | OBF | Xpath (overview | () | | | | | | |
| | Stora | ge media: | , | | | | | | |
| | RAID | , SSD | | | | | | | |
| | Buffe | ring and caching | 5 | | | | | | |

| | Implementation of relational operators: |
|---|--|
| | Implementation algorithms |
| | Cost functions |
| | Query entimization |
| | duely optimization. |
| | Cost-based query optimization |
| | |
| | Transaction processing: |
| | Flat transactions |
| | |
| | Concurrency control and correctness criteria: |
| | Serializability, recoverability, ACA, strictness |
| | Insulation grade |
| | Lock-based scheduling, 2PL |
| | Multiversions to control concurrency |
| | Optimistic scheduling |
| | Logging |
| | internediate statuses (checkpointing) |
| | Current trends in Information Management: |
| | Main memory databases |
| | Column-based data management |
| | NoSQL |
| | |
| 3 | Qualification objectives / learning outcomes |
| | After successfully taking part, the students will be informed about the basics of Information |
| | Management. They will understand the techniques of creating information management systems |
| | and be able to use the respective models, algorithms, and languages to independently use or create |
| | (parts of) information management systems. They will be able to rate the quality of such systems |
| | based on different quality measures. |
| 4 | Prerequisites for participation |
| | Recommended: Successful participation in the lectures "Funktionale und Objektorientierte |
| | Programmierkonzepte" and "Algorithmen und Datenstrukturen" or equivalent knowledge from |
| | other courses of study |
| | |
| 5 | I ype of examination |
| | Subject examination in written form, 90 min. |
| 6 | Prerequisites for Credit Points |
| | Passing the module exam (100%) |
| | |
| 7 | Grading |
| | Standard |
| | |
| | In this lecture, it is possible to take into account accompanying achievements, which – according to |
| | $\frac{9}{2}$ 2012, can load to an improvement in grades of up to 1.0 |
| | z , z_{012} – can lead to an improvement in grades of up to 1.0. |

| 8 | Applicability of the module |
|----|---|
| | B. Sc. Computer Science |
| | B. Sc. Business Informatics |
| | B. Sc. Psychology in IT |
| | Joint B.A. Computer Science |
| | B. Sc. Sports Science and Computer Science |
| | M. Sc. Sports Science and Computer science |
| | Can also be applicable for other programs in the context of interdisciplinary offers. |
| 9 | Literature |
| | Will be announced. Examples are: |
| | Haerder, Rahm, "Datenbanksysteme - Konzepte und Techniken der Implementierung", Springer 1999 |
| | Elmasri, R., Navathe, S. B.: Fundamentals of Database Systems, 3rd. ed., Redwood City, CA: Benjamin/Cummings |
| | Ullman, J. D.: Principles of Database and Knowledge-Base Systems, Vol. 1 Computer Science |
| | |
| 10 | Comment |
| | |

| Modul | la daca | rintion | | | | | | | | |
|------------------------|---|-------------------|--|--|--|---|--|--|--|--|
| Modul | le desc | ription | | | | | | | | |
| Data M | le Aining | and Mach | ine l ea | rning | | | | | | |
| Modul No. 20-00- | 0052 | Credit Po 6 CP | Credit PointsHours requiredSet6 CP180 h1 | | Self 120 | i-study M Dh dเ 1 | | ule tion nester | Recurrence frequency Usually every winter semester | |
| Langua Germa | Language German and English | | | | Per Coc | Person responsibl Coordinators Web | | ble for the module b, Knowledge and Information | | |
| 1 | Cour | ses in the | module | lle | | | | | | |
| | Course No. | | Course | rse title | | Workload (CP) | | Type of lecture | | Contact hours per week |
| | 20-00 0052 |)- -iv | Data N Learni | Mining and Machine ng | | 6 | | Integrat | ed course | 4 |
| | 0052-ivLearningContentDue to the rapid development of informaavailable – and they often contain implicigreat economic or scientific significance.potentially useful knowledge in large setstechnologies in this area.The lecture provides an introduction to thaspect of Data Mining. Methods from diffreferences to exemplary applications. Praoperationalize the knowledge.• Introduction (basic terms, learning prof• Rule learning• Learning of individual rules (generalizatHypothesis rooms, Version Spaces)• Learning of rules-sets (Covering Strateg• Evaluation and cost-sensitive learning (• Instance-based learning (ID3, C4.5, etc.)• Ensemble methods (bias/variance, bag• Pre-processing (Feature Subset Selection | | | | ion te know ata M of dat e field rent p tical e ems, on and , evalu ccura , RISE ng, ra a, disc | chnology, ther ledge that, if it lining is a resea a, and Machine of Machine Le paradigms of N exercises with concepts, exar d specialization uation of rules cy, X-Val, ROC) ndomization, la retization, sam rules (a priori) | e are i t actua arch a e Lear Aachin learnin nples, n, struc , Pruni curves boostin npling, | increasing rea that se ning is one g, with a sp le Learning ng algorith represent ctured ing, multi- s, cost-sen ng, stackin data clear | data volume e known, con eeks to find e of the key oecial focus of g are present ms serve to ation) class probler sitive learnin ng, ECOCs) ning) | es uld be of on the ted with ms) |
| 3 | Clustering and Learning based on assoc Qualification objectives / learning outco After successfully taking part, the studen understand and explain basic techniqu independently use practical Data Minimized meaknesses critically assess new developments in | | | es / learning outcom ng part, the students plain basic technique practical Data Minin v developments in th | nes will k es of E g syst his fie | be able to Data Mining an ems and unde Id | d Mac rstand | hine Learr their stre | ning ngths and | |

| 4 | Prerequisites for participation |
|----|--|
| 5 | Type of examination |
| | Subject examination in written and oral form, 60-120 min. / 30 min. |
| 6 | Prerequisites for Credit Points |
| | Passing the module exam (100%) |
| 7 | Grading Standard |
| | In this lecture, it is possible to take into account accompanying achievements, which – according to § 25 (2) of the 4th amendment of the APB and the rules adopted by the Department 20 on October 2, 2012 – can lead to an improvement in grades of up to 1.0. |
| 8 | Applicability of the module |
| | B. Sc. Computer Science |
| | M. Sc. Computer Science |
| | M. Sc. IT Security |
| | M. Sc. Autonomous Systems |
| | M. Sc. Visual Computing |
| | M. Sc. Internet and Web-based Systems |
| | M. Sc. Business Informatics |
| | B. Sc. Psychology in IT |
| | Joint B.A. Computer Science |
| | B. Sc. Sports Science and Computer Science |
| | M. Sc. Sports Science and Computer Science |
| | B. Sc. Information Systems Engineering |
| | Can also be applicable for other programs in the context of interdisciplinary offers. |
| 9 | Literature |
| | Mitchell: Machine Learning, McGraw-Hill, 1997 |
| | Ian H. Witten and Eibe Frank: Data Mining: Practical Machine Learning Tools and |
| | Techniques with Java Implementations, Morgan-Kaufmann, 1999 |
| 10 | Comment |
| | |

| Module | lule description | | | | | | | | |
|------------|---|------------------|--|-----------------|------------------|-------------------|----------------------------|--------------------------------|----------------|
| Module | - | | | | | | | | |
| Data Mir | ning and N | Machine Learn | ing | | | | | | |
| | | | | | | | | | |
| Module | No. | Credit | Hours required | Self | -study | Mod | ule | Recurrence | e |
| 20-00-0052 | | Points | 180 h | 120 | h | dura | tion | frequency | o.r. (|
| | | 0 (12 | | | | 1 Ser | nester | winter sen | ery hester |
| | | | | | | | | | |
| Languag | е | | | Pers | son responsib | le for | the modul | le | |
| German | and Engli | sh | | Coo | rdinators Web | o, Knov | wledge an | d Informatio | on |
| | | | | Pro | cessing | | | | |
| 1 | Courses | of the Module | 9 | | | | • | | |
| | Course | Course title | | | Workload (C | CP) | Type of I | lecture | Contact |
| | NO. | | | | | | | | hours |
| | | | | | | | | | week |
| | 20-00- | Data Mining | and Machine Learni | ng | 6 | | Integrat | ed course | 4 |
| | 0052-iv | | | | | | | | |
| 2 | Content | | | | | | | | |
| - | Due to t | he rapid devel | opment of informati | on te | chnology, the | re are | increasing | data volum | es |
| | available | e – and they of | ten contain implicit | know | ledge that, if i | t actua | ally becam | , ie known, co | ould be of |
| | great ec | onomic or scie | ntific significance. D | ata N | lining is a rese | arch a | rea that se | eeks to find | |
| | potentia | illy useful know | wledge in large sets c | of dat | a, and Machin | ie Lear | ning is one | e of the key | |
| | teennoie | | | | | | | | |
| | The lect | ure provides a | n introduction to the | e field | of Machine L | earnin | g, with a s | pecial focus | on the |
| | aspect o | f Data Mining. | Methods from diffe | rent | paradigms of I | Machir | ne Learnin | g are preser | nted with |
| | referenc | es to exempla | ry applications. Prac | tical e | exercises with | learni | ng algorith | nms serve to | |
| | operatio | nalize the kho | wiedge. | | | | | | |
| | Introd | uction (basic t | erms, learning probl | ems. | concepts, exa | mples. | represent | tation) | |
| | • Rule le | earning | , 0 , | , | | 1 | -1 | , | |
| | o Learni | ng of individua | al rules (generalizatio | on and | d specializatio | n, stru | ctured | | |
| | Hypothe | esis rooms, Ver | rsion Spaces) | L | | . D | | | |
| | • Evalua | ng of rules-set | s (Covering Strategy, sensitive learning (A | , evai ccura | cv X-Val ROC | s, Prun `curve | ing, muiti- is cost-ser | -class proble sitive learni | erris) ing) |
| | Instance-based learning (kNN, IBL, NEAR, RISE) | | | | | | | | |
| | • Decision tree learning (ID3, C4.5, etc.) | | | | | | | | |
| | Ensemble methods (bias/variance, bagging, randomization, boosting, stacking, ECOCs) | | | | | | | | |
| | Pre-pr Cluster | ocessing (Feat | ure Subset Selection | , disc | retization, sar | npling | , data clea | ning) | |
| | - ciuste | ing and Leaffi | ing based on associa | | uies (a priori) | | | | |
| | | | | | | | | | |
| | | | | | | | | | |

| 3 | Qualification objectives / learning outcomes |
|----|---|
| | After successfully taking part, the students will be able to |
| | • understand and explain basic techniques of Data Mining and Machine Learning |
| | • independently use practical Data Mining systems and understand their strengths and |
| | weaknesses |
| | critically assess new developments in this field |
| 4 | Prerequisites for participation |
| 5 | Type of examination |
| | Subject examination in written and oral form, 60-120 min. / 30 min. |
| 6 | Prerequisites for Credit Points |
| | Passing the module exam (100%) |
| 7 | Grading |
| | Standard |
| | In this lecture, it is possible to take into account accompanying achievements, which – according to |
| | § 25 (2) of the 4th amendment of the APB and the rules adopted by the Department 20 on October |
| | 2, 2012 – can lead to an improvement in grades of up to 1.0. |
| 8 | Applicability of the module |
| | B. Sc. Computer Science |
| | M. Sc. Computer Science |
| | M. Sc. IT Security |
| | M. Sc. Autonomous Systems |
| | M. Sc. Visual Computing |
| | M. Sc. Internet and Web-based Systems |
| | M. Sc. Business Informatics |
| | B. Sc. Psychology in IT |
| | Joint B.A. Computer Science |
| | B. Sc. Sports Science and Computer Science |
| | M. Sc. Sports Science and Computer Science |
| | B. Sc. Information Systems Engineering |
| | Can also be applicable for other programs in the context of interdisciplinary offers. |
| 9 | Literature |
| | Mitchell: Machine Learning, McGraw-Hill, 1997 |
| | Ian H. Witten and Eibe Frank: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations, Morgan-Kaufmann, 1999 |
| 10 | Comment |
| | |
| | |
| | |

| | dule description | | | | | | | | | |
|-------------------------------|---|--|--|--|-------------------|---------------------|----------------------------------|--------------------|---|---------------------------------|
| Modul | e desc | ription | | | | | | | | |
| Genera | e al Com | nutor Sci | onco II | | | | | | | |
| Modu No. 20-00- 0290 | le - | Credit points 6 CP | | Hours required 180 h | Self- 120 | • study h | Module duration 1 Semester | | Recurrence frequency: every other semester | |
| Langua | age | | | 1 | Person respon | | le for the n | nodule | | |
| Germa | n | | | | Vice | Chairperson, | Academic | Affairs | | |
| 1 | L Courses of the m | | e module | | | • | | | | |
| | Cour | se No. | Course | title | | Workload (| CP) | Type of lecture | | Contact hours per week |
| | 20-0 0290 | 0-)-iv | Genera | al Computer Science | II | 0 | | Integ cours | rated se | 4 |
| 2 | Cont In th base * Va * Inh * Arn Adva * Gra * Inp * Err Algo * Re * Sol * Ba * Sea * Sea * Tre | etition of riables, ty rables, ty rays and o anced kno aphical U out / Outy or handli rithms ar cursion rting algo tches, list arches ees and g | , student anced co basic Jav ypes, clas , abstract collection owledge: ser Interf out ing and e ind data st orithms ts, queue raphs | es vill learn the basic ncepts of the Java p va-knowledge: sses, program flow t classes, interfaces faces faces xceptions tructures: | algorit rogram | thms and data | e. | s of Cor | nputer Sci | ience, |
| 3 | Qualification objectives / learning outcomes After attending the event, students will be able to develop extensive programs in Java to independently use the basic algorithms and data structures of Computer Science to estimate the advantages and disadvantages of elementary algorithms with regard to complete and execution time | | | | | | | complexity | | |
| 4 | Prere Recc Gene - bas - bas - exp | equisites ommende eral Infor ic progra ic knowle perience i | for partie ed: matics I c mming s edge in C n workin | cipation or kills omputer Science g with computers | | | | | | |

| 5 | Type of examination |
|----|---|
| | Module-accompanying examination: |
| | |
| | * [20-00-0290-iv] (Subject examination in written and oral form, Standard) |
| 6 | Prerequisites for Credit Points |
| | Passing the module exam (100%) |
| 7 | Grading |
| | Module-accompanying examination |
| | * [20-00-0290-iv] (Subject examination in written and oral form, weighting: 100%) |
| | |
| 8 | Applicability of the module |
| | B. Sc. Computer Science |
| | M. Sc. Computer Science |
| | Can also be applicable for other courses of study |
| | |
| 9 | Literature |
| | Java lernen mit BlueJ: Eine Einführung in die objektorientierte Programmierung David J. Barnes, |
| | Michael Kölling Pearson Studium 4., aktualisierte Auflage, 2009 |
| | ISBN-13: 978-3-8689-4001-5 |
| | Algorithmen in Java |
| | Robert Sedgewick |
| | Pearson Studium |
| | 3. überarbeitete Auflage, 2003 |
| | ISBN-13: 978-3-8273-7072-3 |
| | Einführung in die Programmierung mit Java Robert Sedgewick, Kevin Wayne Pearson Studium 1. |
| | Auflage, 2011 |
| | ISBN-13: 978-3-8689-4076-3 |
| 10 | Comment |
| | |

| Module | e descr | ription | | | | | | | | | | | |
|----------------------------------|--------------------------------|---|---|--|--|--|------------|-----------------------------|-------------|-------------------------------|--|--|--|
| Module | 9 | - | | | | | | | | | | | |
| Genera Modul 20-00- | l Comp e No. 0304 | outer Scie Credit Po 6 CP | ence l oints | Hours required 180 h | Self-stu 150 h | Self-study Modul 150 h durati | | ule Recurre tion frequer | | ce y : every | | | |
| Langua | Øe | | | | Person | responsible | for the m | odule | Utilet sen | Ilestei | | | |
| Germai | b- n | | | | Vice Chairperson, Academic Affairs | | | | | | | | |
| 1 | Cours | ses of the | e module | | I | • | | | | | | | |
| | Cours | se No. | Course t | itle | V | Vorkload (C | CP) | Type of lecture | | Contact hours per week | | | |
| | 20-00 | -0304-iv | General | Computer Science I | | 0 | | Integr course | ated e | 2 | | | |
| 2 | Cont | ent | | | | | | I | | | | | |
| 3 | • • • • The la | Introdu Binary Elemer Von Ne Elemer Basics o ecture wi ification o | uction to numeric htary logi eumann c htary con of compu Il be acco objective Basic kn Practica Basic pr | programming (Kare and information rej cal and arithmetic of computer architectu cepts of operating s iter networks ompanied by progra s / learning outcom lowledge concernin l use of computers ogramming skills | ell, Java, or presentati operations ure systems mmming ex nes g the mos | ⁻ similar) on ercises. t important | concepts c | of Com | puter Scier | nce | | | |
| 4 | Prere | quisites | for partio | cipation | | | | | | | | | |
| 5 | Type Modu * [20- | of exam Ile-accom 00-0304- | ipanying iv] (Subje | examination ect examination in w | vritten and | l oral form, | Standard) | | | | | | |
| | | |] () - | | | · · · · , | , | | | | | | |
| 6 | Prere Passir | quisites f ng the mo | or Credit | Points m (100%) | | | | | | | | | |
| 7 | Gradi Modu | ng Ile-accom | ipanying | examination | | | | | | | | | |
| | * [20- | 00-0304-i | iv] (Subje | ct examination in w | vritten and | l oral form, | weighting: | 100% | | | | | |

| 8 | Applicability of the module |
|----|--|
| | B. Sc. Computer Science |
| | M. Sc. Computer Science |
| | Can also be applicable for other courses of study |
| | |
| 9 | Literature |
| | Literatur: David J. Barnes und Michael Kölling, Java lernen mit BlueJ: Eine Einführung in die |
| | objektorientierte Programmierung, Pearson Studium; 5te Auflage (1. Januar 2013), ISBN 3868949070 |
| | |
| 10 | Comment |
| | |

| Mod | lule desc | ript | ion | | | | | | | | |
|--------------------|------------------------|---|---|---|--|--|----------------------------------|--------------------------------|--|---------------------------------|--|
| Mod | lule | | | | | | | | | | |
| Four | ndations | of L | anguage Tech | nnology | | | | | | | |
| Mod No. 20-0 | l ule 0-0546 | Cre 6 (| edit Points CP | Hours required 180 h | Self-study 120 h | | Moo dura 1 Se | Jule ation mester | Recurrence frequency: usually every winter semester | | |
| Language German | | | | | Person responsible for the module Coordinators Web, Knowledge and Information Processing | | | | | | |
| 1 | Courses of the Module | | | | | | | | | | |
| | Cour No. | Course Course title No. | | 2 | | Workload | I (CP) | Type of | lecture | Contact hours per week | |
| | 20-00 0546 | 0- -iv | Foundation: Technology | s of Language | | 6 | | Integrated course | | 4 | |
| | The I auto | ectu mati copic b T b S b S b C b S b C b I l b E c S b V b V b C C c S c S c S c S c S c S c S c S c S c S | ire offers an i ic language tech okenization ogmentation opeech recogn corpora tatistical ana Machine learn Categorization nformation eintroduction t Data structure tructured pro Vorking with Use of librarie program libra | ntroduction to the k echnology – based o nology / Natural lan nition lysis n and classification xtraction o Python es ogramming files s ry NLTK | key poir on the e | nts of view xample of processing | , probler the prog g (NLP) | ns, metho ramming l | ds, and tech anguage Pyt | niques of hon. | |
| | addr addr | ectu ess t iring | he is based o he theoretica extensive pr | al methods in an exp ogramming knowled | blorator olorator dge. | rython – V y and prob | olem-orie | ented man | iner, withou | t | |

| 3 | Qualification objectives / learning outcomes |
|----|--|
| 3 | Qualification objectives / learning outcomes |
| | After attending the lecture, the students will be able to |
| | define the basic terminology of automatic language technology, |
| | name and explain essential questions in this area |
| | explain and implement simple Python programs |
| | transfer the discussed methods and techniques to concrete application scenarios of text |
| | comprehension |
| | critically assess their knowledge and their limits. |
| 4 | Prerequisites for participation |
| | Recommended: |
| 5 | Type of exam |
| 5 | Module-accompanying examination |
| | |
| | (Subject examination in written and oral form, Standard) |
| 6 | Prerequisites for Credit Points |
| Ŭ | Passing the module exam (100%) |
| | |
| 7 | Grading |
| | Standard |
| 0 | |
| 8 | Applicability of the module |
| | B. Sc. Computer Science |
| | M. Sc. Computer Science |
| | M. Sc. IT Security |
| | M. Sc. Internet and Web-based Systems |
| | M. Sc. Business Informatics |
| | B. Sc. Psychology in IT |
| | Joint B.A. Computer Science |
| | B. Sc. Sports Science and Computer Science |
| | M. Sc. Sports Science and Computer Science |
| | Can also be applicable for other programs in the context of interdisciplinary offers. |
| | |
| 9 | Literature |
| | Steven Bird, Ewan Klein, Edward Loper: Natural Language Processing with Python, O'Reilly, 2009. ISBN: 978-0596516499. <u>http://www.nltk.org/book/</u> |
| 10 | Comment |
| | |