International Master in Data Analytics

Welcome to Summer Term 2024

Prof. Dr. Dr. Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL)
University of Hildesheim, Germany

April 3, 2024

Outline

1. About us

2. The Goals of the Data Analytics Program

3. The Structure of the Data Analytics Program
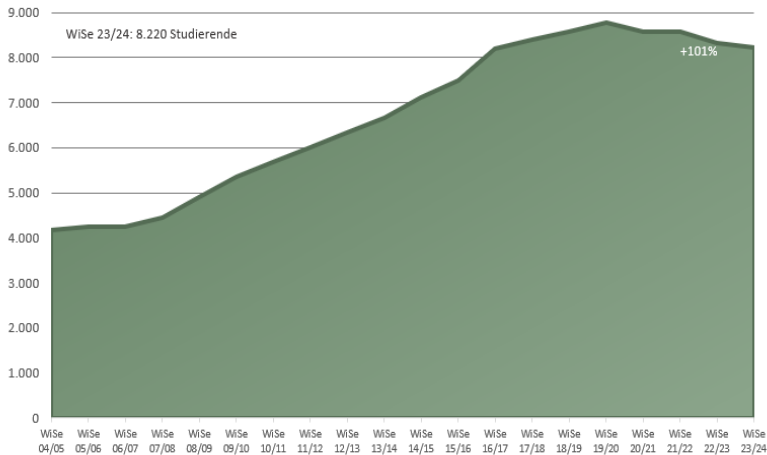
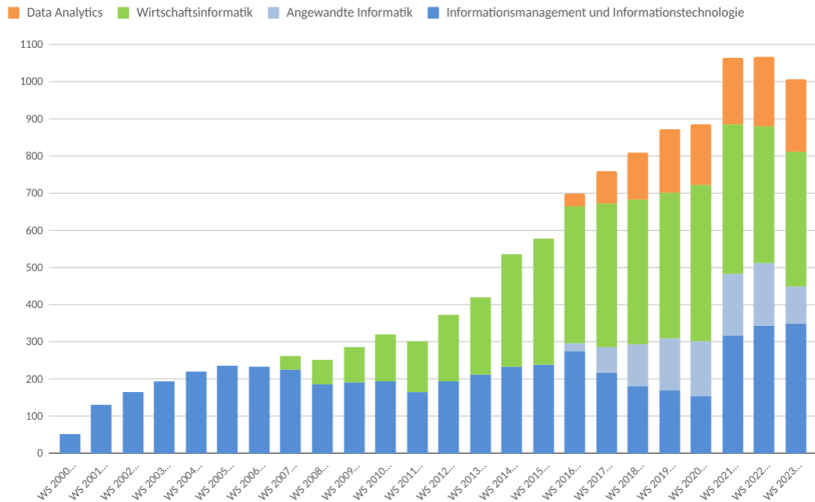4. Exams and Preparation

Outline

University of Hildesheim

▶ small German research university
  ▶ 8220 students
  ▶ 95 professors, 907 employees in research and administration

▶ with focus on
  1. Educational Sciences,
  2. Cultural Sciences and
  3. Computer Science.

▶ in the heart of Germany

Students

Entwicklung der Studierenden (mit Beurlaubten)

Computer Science Students

4 Main University Sites

4 Faculties

1. Educational and Social Sciences
2. Culture Studies and Aesthetic Communication
3. Linguistics and Information Sciences
4. Mathematics, Natural Sciences, Economics and Computer Science

4 Faculties

1. Educational and Social Sciences
2. Culture Studies and Aesthetic Communication
3. Linguistics and Information Sciences
   - ▶ Institute for Information Sciences and Linguistic Technologies
   - ▶ . . .
4. Mathematics, Natural Sciences, Economics and Computer Science
   - ▶ Institute for Computer Science
   - ▶ Institute for Economics and Information Systems
   - ▶ Institute for Mathematics and Applied Computer Science
   - ▶ . . .

4 Faculties

1. Educational and Social Sciences
2. Culture Studies and Aesthetic Communication
3. Linguistics and Information Sciences
   ▶ Institute for Information Sciences and Linguistic Technologies
   ▶ . . .
4. Mathematics, Natural Sciences, Economics and Computer Science
   ▶ Institute for Computer Science
      ▶ Machine Learning (ISMLL) — Prof. Schmidt-Thieme
      ▶ Data Science — Prof. Landwehr
      ▶ Software Engineering — Prof. Schmid
      ▶ Intelligent Information Systems — Prof. Althoff
   ▶ Institute for Economics and Information Systems
   ▶ Institute for Mathematics and Applied Computer Science
   ▶ . . .

ISMLL

- ▶ Information Systems and Machine Learning Lab
- ▶ research group focused on
    - ▶ **supervised machine learning**
    - ▶ for complex data and
    - ▶ complex decisions

- ▶ professor, postdoc & 15–20 PhDs
- ▶ over 200 papers,
  many at the best Machine Learning conferences and journals
- ▶ several best paper awards
- ▶ won ECML challenge 2009 and 2016
- ▶ 4 multi million Euro European research projects with industry partners
- ▶ many focused research projects
- ▶ 8 professors emerged from the group within the last 15 years

VWFS DARC

- ▶ **V**olkswagen **F**inancial **S**ervices **D**ata **A**nalytics **R**esearch **C**enter
- ▶ Recent establishment of a research center, focused on
  - ▶ Advanced Machine Learning Concepts
    - ▶ Deep time series forecasting
    - ▶ Computer vision
    - ▶ Recommender systems
    - ▶ Reinforcement learning
  - ▶ Trustworthy AI
    - ▶ Uncertainty quantification
    - ▶ Explainable AI
    - ▶ Transfer and Multi-task learning
    - ▶ Data & ML model monitoring
- ▶ professor, 2 postdocs & 8-10 PhDs
- ▶ collaboration in its 6th year
- ▶ over 10 papers, many at the best Machine Learning conferences
- ▶ biggest industry partner, working on interesting ML problems
- ▶ many focused research projects

Outline

Goals of the Program

▶ a deep and thorough introduction to cutting edge research in
  ▶ Machine learning,
  ▶ Big Data and
  ▶ analytical technology

▶ complementary training in selected application domains
  ▶ marketing, logistics, computer science, environmental science

▶ brings together students from all over the world and different background disciplines
  ▶ completely taught in English

▶ Data Analytics is a **research Master program**.

Program Requirements

- ▶ DA targets students with an analytical Bachelor's Degree
  - ▶ Computer Science, Information Technology
  - ▶ Mathematics, Statistics
  - ▶ Business Administration, Economics
  - ▶ and related fields

- ▶ Required proficiencies:
  - ▶ math
  - ▶ programming
  - ▶ English

Data Analytics Students

Intake	Applied	Enrolled	Countries
SoSe 2024	2852	23	11
WiSe 2023/24	2647	34	18
SoSe 2023	1909	21	16
WiSe 2022/23	2343	40	18
SoSe 2022	1348	26	16
WiSe 2021/22	2116	26	15
SoSe 2021	1301	35	20
WiSe 2020/21	1798	26	12
SoSe 2020	1488	19	10
WiSe 2019/20	2122	31	15
SoSe 2019	1407	47	20
WiSe 2018/19	1896	26	14
SoSe 2018	1116	25	15
WiSe 2017/18	1012	39	18
SoSe 2017	470	27	13
WiSe 2016/17	170	31	13
Total	23986	443	
Alumni		240	

Outline

Program Structure

▶ 4 semesters spanning over two years

▶ total 120 CPs (credit points) which are divided into
  ▶ A methodological core (65%)
  ▶ An application area (10%)
  ▶ A master's thesis (25%)

Courses First Year (Summer Start)

	Nr	Module	Type	CPs
1st Term	M5	Big Data Analytics	Lecture	6
	M7	Data and Privacy Protection	Seminar	3
	M8	Distributed Data Analytics	Lab	6
	MS1	*Methodological Specialization*	Lecture	6
	AM1	*Application Module I*	-	6

	Nr	Module	Type	CPs
2nd Term	M1	Machine Learning	Lecture	6
	M2	Modern Optimization Techniques	Lecture	6
	M3	Programming Machine Learning	Lab	6
	M4	Seminar Data Analytics I	Seminar	4
	M10	Project (part I)	Project	3
	AM2	*Application Module II*	-	6

Courses Second Year (Summer Start)

	Nr	Module	Type	CPs
3rd Term	M6	Advanced Machine Learning	Lecture	6
	M9	Seminar Data Analytics II	Seminar	4
	M10	Project (part II)	Project	9
	M14	Master Thesis (part I)	Project	6

	Nr	Module	Type	CPs
4th Term	M11	Planning and Optimal Control	Lecture	6
	M12	Project (part III)	Project	3
	M13	Seminar Data Analytics III	Seminar	4
	M14	Master Thesis (part II)	Project	24

▶ as the 4th term is 7 CPs too heavy,
  better plan for a 5th term (at least for 2 more months).

Courses First Year (Summer start, with prior ML background)

Nr	Module	Type	CPs
M5	Big Data Analytics	Lecture	6
M7	Data and Privacy Protection	Seminar	3
M8	Distributed Data Analytics	Lab	6
M4	Seminar Data Analytics I	Seminar	4
MS1	*Methodological Specialization*	Lecture	6
AM1	*Application Module I*	-	6

1st Term

Nr	Module	Type	CPs
M1	Machine Learning	Lecture	6
M2	Modern Optimization Techniques	Lecture	6
M3	Programming Machine Learning	Lab	6
AM2	*Application Module II*	-	6
M9	Seminar Data Analytics II	Seminar	4
M10	Project (part I)	Project	3

2nd Term

Courses Second Year (Summer Start, with prior ML background)

	Nr	Module	Type	CPs
3rd Term	M6	Advanced Machine Learning	Lecture	6
	M13	Seminar Data Analytics III	Seminar	4
	M12	Project (part II)	Project	9
	M14	Master Thesis (part I)	Project	6

	Nr	Module	Type	CPs
4th Term	M11	Planning and Optimal Control	Lecture	6
	M12	Project (part III)	Project	3
	M14	Master Thesis (part II)	Project	24

Elective Courses 1: Methodological Specialization

▶ to deepen your methodological understanding and
   widen the models and methods you command.

▶ currently 5+3 courses:
   ▶ Machine Learning for IT Security (Landwehr) – Summer
   ▶ Advanced Computer Vision (Landwehr) – Summer
   ▶ Advanced Case Based Reasoning (Althoff) – Summer
   ▶ Time Series Analysis (Mentemeyer) — Summer
   ▶ Large Language Models (Stubbemann) — Summer

   Courses paused in Winter 2022/23 and Summer 2023:
   ▶ Bayesian Networks — every odd Summer: 2024
   ▶ Computer Vision — every even Summer: 2025
   ▶ Business Analytics — Winter

▶ you have to choose at least one course (6 CP)
   ▶ marks of just one course count to your final degree

Elective Courses 2: Application

▶ to provide a testbed for applying data analytics methods.

▶ currently from 6 areas:
  ▶ Computer Science / Software Engineering
  ▶ Computer Science / Media Systems
  ▶ Business Administration
  ▶ Information Retrieval and Information Sciences
  ▶ Natural Language Processing
  ▶ Environmental Sciences

▶ you have to choose courses worth at least 12 CP from one area
  ▶ e.g., two lectures with tutorials
  ▶ marks of courses worth 12 CP count to your degree

Student Research Projects

▶ to provide a testbed for applying data analytics methods.

▶ currently structured with 15 CPs over 2 terms:
  ▶ Work closely under supervision in teams
  ▶ Students present final outcomes in SRP Conference

▶ several state-of-the-art papers published
  ▶ German AI Conference (2021, 2022)
  ▶ DASC (2023)
  ▶ ECDA (2019, 2020)
  ▶ ECML-Workshops (2018, 2020)
  ▶ DEXA (2019, 2020)
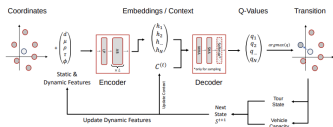

Fig. 1. Architecture overview

Outline

Take Your Studies Seriously

1. Attend the lectures!
2. Take notes in the lectures!
3. Solve the tutorial and lab problems on your own!
4. Read the books!

Example efforts: 2h lecture plus 2h tutorial

- ▶ 6 CP = 180h student effort
- ▶ 4h/week face-to-face
- ▶ 6h/week solving tutorials
- ▶ 2h/week post-preparation and reading
- ▶ 12h exam preparation
- ▶ $(4 + 6 + 2)h/w \cdot 14w + 12h = 180h$

Exam Regulations (1/2)

▶ **examination periods:**
  ▶ exams in the first 4 weeks after the lecture period ends
  ▶ Summer 24: 15.7.2024 – 11.8.2024

▶ **trials:**
  ▶ you have 3 trials for each written exam
    ▶ one at the end of term
    ▶ one at the beginning of the next term
    ▶ one at the end of term next year
      — this usually will prolong your studies
  ▶ we may switch later trials to oral exams

▶ **exam conditions:**
  ▶ may vary from course to course as documented in the course catalogue
  ▶ for your courses at ISMLL:
    ▶ open book

Exam Regulations (2/2)

▶ **duration of studies:**
  ▶ default duration are 4 terms
  ▶ you are welcome to extend by a term or two if you need it
  ▶ after 10 terms you will have to pay long-term study fees
    (or after 16 terms minus the terms needed for your bachelor at a
    German university)

▶ **formal regulations:**
  ▶ Masterprüfungsordnung Informationsmanagement und
    Informationstechnologie (currently in German only) and
  ▶ Course Catalogue International Master in Data Analytics

What to get Done Before Your Studies Start (1/2)

▶ an account at our computing center
  ▶ will allow you to register for courses
  ▶ account information has been sent by electronic mail

▶ register for all your courses at the teaching information system LSF
  ▶ LSF: Lehre–Studium–Forschung: Teaching–Studies–Research
  ▶ Machine Learning, Programming Machine Learning Lab, Modern Optimization Techniques
  ▶ a specialization and an application course

▶ get a computer/laptop you can work on whenever you have to
  ▶ programming editor/IDE
  ▶ compiler/interpreter (esp. Python)
  ▶ programming language documentation
  ▶ LaTeX (or OpenOffice)

What to get Done Before Your Studies Start (2/2)

- ▶ get your first books
  - ▶ at least Murphy

- ▶ refresh your Math
  - ▶ at least Murphy, ch. 2
  - ▶ Murphy, ch. 1–6

- ▶ refresh your programming skills
  - ▶ esp. Python

- ▶ find a quiet place to work

Whom to ask

Questions or issues regarding	ask
exercises	course tutor
lecture	course lecturer
program	program director
program director	study dean
computers	computing center, room E114

Welcome to University of Hildesheim!

Welcome to the
International Master in Data Analytics!

I wish you successful studies!