Evaluation of the Applied Data Science program at the University of Gothenburg

2021-04-13

1 Introduction

This document presents a summary of the evaluation of the Master's program in Applied Data Science at the University of Gothenburg performed in Autumn 2020 and Spring 2021 by an external evaluation group. It starts with an introduction, continues with the evaluation section according to the predefined criteria and ends with a summary.

1.1 Background

The IT-faculty board at the University of Gothenburg (GU) decided to recruit four members to evaluate its international Applied Data Science Master's Program (N2ADS)[1] hosted at the Department of Computer Science and Engineering that is a joint department at the University of Gothenburg and Chalmers University of Technology. This decision was taken on 30-09-2020, see faculty document reg no. 2020/1586. The members that were appointed to the evaluation group are: professor Patrick Lambrix (chair), department of computer and information science at Linköping University, fil. doctor Dana Dannélls, Språkbanken Text, department of Swedish, University of Gothenburg, fil. doctor Mattias Grönkvist, Jeppesen and student representative, Emil Bagge, Kings College London.

On November 13, 2020 the evaluation group had together with the faculty and the department representatives an inauguration meeting to discuss and plan for the evaluation process. Prior to the meeting a large amount of written documents were made available to the group. These included materials provided by the IT faculty about the program, the program evaluation process, the program syllabus and local qualification descriptor, the course literature lists, evaluation material and examination documents. The evaluation group has studied the material individually and had two joint online meetings in January to prepare for a site visit with the program leader, teachers, students, alumni and administrators. The online site visit took place on 2021-02-04 via Zoom. The program for the site visit is available in Appendix A.

[1] https://www.gu.se/en/study-gothenburg/applied-data-science-masters-programme-n2ads

1.2 The program content and structure

The international Applied Data Science Master's Program, a two-year master's program, encompasses a total of 120HP, and started in the Autumn of 2017. Today it consists of courses in Data Science and related subjects. Eligible to apply for the program are students with an undergraduate degree with at least three years full-time study in diverse topics such as cognitive science, software engineering, information technology, and bioinformatics.

The program syllabus and local qualification descriptor have undergone a series of changes since program start. The qualification descriptor for the program Data Science at the University of Gothenburg was confirmed by the IT Faculty board on 2017-09-13 (reg. no. G 2017/296) and revised several times. The latest version is from 2020-04-08. In the first version of the program syllabus and the local qualification descriptor, the number of courses in the main field of study Data Science were few. As more courses in the main field of study were created, they were also added to the local qualification descriptor. According the latest syllabus, the program comprises nine compulsory courses including the master's thesis course. Further, the program offers a number of elective courses covering specific topics. The list of courses currently offered by the program (and for which course material was made available to the evaluation group) is available in Appendix B.

The program is popular with about 700 applicants each year for 50 positions. The program normally accepts 65 students of which around 40 actually start the program.

1.3 About the evaluation criteria

The program evaluation follows the university's criteria according to the policy for quality assurance and continuous quality improvement and education at the University of Gothenburg.[2] More specifically, the assessment followed the following 10 criteria: (1) Goal fulfillment; (2) Student-centered learning; (3) Education's scientific and experience base; (4) Teacher competence and capacity; (5) Relevance of education for students and society; (6) Student influence; (7) Accessibility and effectiveness of learning environment; (8) Continuous evaluation and development of the education program; (9) Conformance to laws and regulations; (10) Sustainability and equal opportunity. Further, we add information about some other observations that do not fit within the above listed ten evaluation criteria.

[2] `https://medarbetarportalen.gu.se/digitalAssets/1708/1708934_utbildningsutvardering-med-extern-bedomning-vid-uf.pdf`

2 Evaluation following the criteria

2.1 Goal fulfillment

The evaluation group has analyzed the program syllabus and local qualification descriptors, and can conclude that they fulfill the goals and the requirements set for the program. The evaluation group's overall impression is that the Applied Data Science master program is of good quality.

The program syllabus is normally presented to the students by the program leader in the beginning of the first year in a joint session. During and after this session the students can pose follow-up questions. To make the structure of the program more transparent, the evaluation group recommends to include a matrix of the program syllabus including information on dependencies between courses on the program web page. In this way it is easier for students to achieve an overview of the content of the program and they can prepare their choices for elective courses in a better way.

The program makes a good effort to ensure that the 'Applied' in the program's name is in focus. For example, real-world problems and domains are considered in the courses, there is a focus on applying methods rather than covering the theory behind them, many courses have guest lecturers from industry, and many students have the opportunity to work with companies during projects and master's thesis work. The application side of the program is based on theoretical foundations and many courses contain both theory and practice. All the groups we interviewed agree there is a good balance between theory and practice. One issue that has been pointed out to the evaluation group multiple times is that not all students in the program have a strong technical background, which makes it difficult to adapt some courses to the right technical level. The teachers in the program are aware of this and take action in this respect. One possible way to improve the balance between theory and practice further could be to organize small workshops where several people from the industry and the academy work on solving a particular problem together.

All courses are evaluated at the end. The course evaluation reports for each of the compulsory courses have been made accessible to the evaluation group. After reviewing these, no noteworthy issues have been brought to the evaluation group's attention. The response rate of the students seems adequate.

While there is a consistent evaluation procedure for the individual courses, an evaluation of the program as a whole is lacking. The evaluation group recommends the design of an evaluation protocol for the whole program.

There is no systematic follow-up of the courses with respect to the program as a whole. The evaluation group sees a need for a more systematic control over the program courses, that is how many students are registered for the courses, how many complete the courses and program. Having a systematic follow up strategy will help to learn the program success rate over time. Regarding the goals on program level, the evaluation group has looked at the different courses and decided that the goals are most likely reached by the current set of courses. However, it is not clear that there is a systematic control that the program goals

are achieved by all students regardless of which elective courses they choose. The evaluation group recommends maintaining a matrix with the list of all the courses included in the program on one axis and the list of program goals on the other axis. For each course, there should be a specification of which program goals have been fulfilled. This will help to keep track of goal fulfillment on the program level. Further, information on course level regarding goal fulfillment as well as dependencies between courses is important. This will help visualize progression which is especially important when courses and program content change.

The evaluation group received statistics of the participation of students in the program but unfortunately, as the program leader also points out, there is no way to easily separate the Applied Data Science students from the group of all GU students. However, the evaluation group considers it important to obtain such information for evaluation and further development purposes. From the interviews it follows that completion rates for courses are normal. About one fourth of the students who start the program successfully complete it.

The examples of Master's thesis reports included in the evaluation repository maintain a good national level. About one third of the Master's projects are carried out with industrial partners. The majority of Master's thesis projects have been so far completed successfully.

2.2 Student-centered learning

There are different definitions of student-centered learning ranging from putting responsibility for the learning on the students while teaching skills on how to learn a subject, to acknowledging that each student is an own individual with their own abilities and interests where the teacher becomes a facilitator in the learning process. For the purpose of this evaluation a broad interpretation of the term has been used.

The course plans for the different courses clearly state learning outcomes in terms of three categories: knowledge and understanding, competence and skills as well as judgement and approach. This is a basis for students to be able to take responsibility in the learning.

There are several project courses as well as mini-projects in other courses and the master's thesis that allow students to gear the learning process for knowledge and skills towards their own interests. Group projects also enable learning of soft skills.

Some courses used the flipped class room method that is one of the methods promoting student-centered learning. Further, the fact that lecturers recorded their lectures and made them available allows the students to choose when to study and at which pace.

2.3 Education's scientific and experience base

The content and form of the teaching rests on a scientific basis and proven experience. A large majority of the teaching staff are leading researchers in

their fields, and the connection to current research is strong.

The course literature covers both theory, methods and applications and consists mostly of articles published in international conferences and course books. The content of courses included in the program is adapted to state of the art approaches. For almost all courses the literature is freely available online.

The content of the courses conveys the whole pipeline of data gathering, processing and analysis, something which contributes to a better understanding of the field of Data Science. Elective courses related to the program such as computer and network security have contributed to preparing the students for the industry. However, the alumni certify that more mathematics (e.g., linear algebra) is needed in the beginning of the program, as well as a course about research methods in data science, learning how to apply them more systematically. More practice with visualisation tools should be incorporated within the different courses as well.

The content of the program is undergoing continuous changes. The evaluation group could observe that over the last years there have been some minor changes in the course content of some courses. For example, the course requirements for *Introduction to Data science* and *Techniques for large-scale data* seem insufficient because many students are having a hard time catching up. Both teachers and the program leader are aware of these difficulties and have acknowledged there is a need for improvement.

Not all of the courses have connections with industrial partners. In the courses where industrial partners are involved there is a fruitful collaboration. Invited industry speakers are in many cases interacting with the program teachers in one way or another and have a good insight into the program content. Industrial partners are often alumni, some have research collaboration with teachers, some are adjunct faculty members.

With regard to scientific best practice, an increasingly important issue for data scientists is data management (software engineering methods / project management) with data science focus which is currently missing. The evaluation group's suggestion is to include these topics in existing courses or as individual courses.

2.4 Teacher competence and capacity

It is the opinion of the evaluation group that the teachers of the program have the pedagogical and subject matter competence needed to fulfill the program needs. The teachers have basic pedagogical training and students are in general satisfied with the pedagogical aspects of the courses. The students interviewed by the advisory group were of the opinion that virtual lectures over Zoom had worked well during the pandemic, and remarked that some teachers had used creative new solutions for a number of the virtual lectures.

Many courses in the program have guest lecturers from companies and institutions using Data Science. These lecturers are not required to have any pedagogical training. However, this is not considered a problem since the guest lectures are not integral parts of the course curricula, but rather have the pur-

pose of broadening the students' perspectives with industrial uses of Data Science.

The teachers in the program teach about 50% of their working time. They express that their teaching load is balanced and not a cause for concern. However, it was noted that the large number of topics for Master's thesis projects has started to make it difficult for the teachers to have time to supervise all the projects. As the number of students increases, there is concern that this might become an even more serious problem. The teachers also remarked that the initial development of the program, as well as changes to courses over time temporarily increases the workload.

Teaching assistant levels are organized centrally with minor local variations depending on course loads. Neither students nor program representatives have expressed that teaching assistant availability is a problem in the program.

Given the concerns expressed by both teachers and program leaders about the increasing deficit in teaching resources, the evaluation group recommends that the teaching resource need for the program is investigated and potentially increased.

As Data Science is a fairly new research field in its current form, a number of the teachers do not have their academic background in Data Science but rather in neighboring fields such as Computer Science. The evaluation group has not found this to be a cause for concern. The teachers who initially developed the program have a solid research focus on Data Science, and the teachers with a different research background have in many cases shifted their research focus towards Data Science.

Further, the formation of the Data Science and AI division within the Computer Science and Engineering department has put Data Science, Artificial Intelligence and Machine Learning in focus. This has lead to a number of department recruitments over the last few years who specialize in Data Science. As a further indication of the high subject matter competence of the teachers in the program, the Data Science and AI division runs many research projects supported by CHAIR (Chalmers AI research initiative) and WASP (Wallenberg AI, Autonomous Systems and Software Program) with participation of the program teachers.

2.5 Relevance of education for students and society

It is the opinion of the evaluation group that the Applied Data Science program as a whole is highly relevant for the needs of society and students.

Data Science has been a very hot topic in industry, academic research and among the general public in the last years, so there is little doubt that the Applied Data Science program fills a need in society by educating more people in this area. Given the current technological trends of gathering and analyzing more and more data in all sorts of applications and hardware, it is reasonable to expect that the need for Data Science competence will even increase in the future.

The students and alumni who have been interviewed by the evaluation group have expressed that they are satisfied with how the program has and is preparing them for their future work life, and the course evaluations have not given the impression that the students think some significant area is missing in the program. There are some suggestions, however, for topics or courses that should be added to the program. These include courses related to mathematics such as statistics, linear algebra and calculus. The existing course on linear statistics methods was mentioned to be really good, but it is not a mandatory course. Further, the students and alumni requested topics such as visualization, handling of large data sets (which is briefly covered at the moment) and research methods in Data Science. Also project management was requested including development methods (e.g. Scrum, Agile) and learning how to work with real data in real life and evaluating input from clients. The evaluation group acknowledges that some of these improvement areas are already partially covered in the program, and some have already been addressed in the last years. However, it is the recommendation of the evaluation group that the program leaders investigate the possibilities to include more elements from the requested areas as a way to further improve the practical usefulness of the program.

The students have described how many courses contain very up-to-date information about the latest state of the art, for example when it comes to machine learning. The alumni mentioned that in a fast moving field such as Data Science you cannot always expect to be up to date with all knowledge, tools, etc, but that the program has provided the necessary skills to assess new tools and processes in a sound way. Our impression is that the responsibility to keep the program up to date with the latest development lies mainly on the teachers in the courses, there is no coordinated effort on a program level to survey current state of the art, or to adjust the program based on the needs of the industry and society. The evaluation group recommends that the program forms some sort of advisory board with industry representatives, as an organized way to ensure that the program caters to the Data Science need of the industry.

Many courses in the program contain guest lectures given by people from the industry or from other external partners, and these lectures were appreciated by the students. They liked how the lectures showed how Data Science is used in practice, but they also appreciated the possibilities to make connections with industry representatives. The students mentioned that some of the guest lectures stuck with them while others were less memorable. As a way to make the guest lectures even more appreciated, while at the same time increasing exposure to real data and problems as requested by the students, the evaluation group recommends that the program leaders consider the possibility of including material from the guest lectures in assignments or projects in courses when possible.

2.6 Student influence

On course level individual students can comment on the courses through course evaluations. There is a mid-course evaluation for changes that can be made

during the current running of the course and end-course evaluations for issues for the next time the course runs. The end-course evaluations follow a standard form for the GU specific courses, while Chalmers takes care of the evaluation of the courses that are common for GU and Chalmers. Both students and alumni stated that they feel that teachers want to receive comments and that the students' comments are taken into account and actions are taken. In some courses changes based on the students' comments could be better communicated. The evaluation group also noticed that the participation in the course evaluations is not very high and there is not always that much actionable information. This is not a program-specific problem, but is known for many programs and universities. Further, each course has student representatives that have a direct contact with the teachers and issues can be taken up and discussed with the course leadership in this way.

Although students can contact the program leader and when this happens, the students feel that issues are dealt with, there is no student representation on program level. This hampers students to take up program level issues with the program leadership. This is a disadvantage for both students and program. The evaluation group considers student representation at the program level as an action point. According to our information a program board may be created and student representatives could be members of such a board.

2.7 Accessibility and effectiveness of learning environment

Due to the covid-19 situation, it was not possible to visit the campus and look at the physical work environment of the students and teachers. During this situation most lectures were given via Zoom and the students commented that these worked fine. Most teachers recorded their Zoom lectures and made them available. Some teachers used flipped class room. The lab rooms were not so accessible to everyone because of an issue with pass cards. The students remarked that it was important to have a good own labtop. The guides for installing and using software used in the courses were on a good level. If more project or laboratory work regarding big data would be included in the program, there may be a need for larger computing capacity that is made available to the students.

Students with disabilities can apply for pedagogical support centrally. When approved, the program and the teachers take this into account and provide special aid for the student.

The communication between the students and the study administration of the program happens via the student portal that, among others, contains a page for the department that is regularly updated. Further, there is much e-mail and telephone contact. During covid-19 times the student reception is open via Zoom four times a week. A recommendation of the evaluation group is that the program organizes follow-up meetings (e.g. lunch meetings) to get to know how the students are doing in different aspects.

2.8 Continuous evaluation and development of the education program

The program is still young and changes have been made for several courses during these first years. More changes are ongoing and the expectation is that this will lead to a stabilization of the program regarding structure, content and personnel in a year. Further efforts are planned after the stabilization regarding quality, pedagogical development and new trends. When the program was developed there were no guidelines for what should be included in such a program. Currently, such guidelines[3] exist and the evaluation group recommends to align to or take inspiration for changes from such guidelines.

One of the challenges for the program is that it is a multi-disciplinary program that requires expertise in different areas. Therefore, there is a need for recruiting more personnel in different competence areas. The current set-up requires a good cooperation with Chalmers as well as temporary assignments.

Further, because of the breadth of the area there may be a need for more tracks that students can follow within the program. At this moment there are essentially two: machine learning and natural language processing. An interesting possible way forward is to find synergy effects with other education programs, in particular the education programs of areas where data science is applied such as biology, astronomy and marine science. The evaluation group encourages this and from our understanding initial contacts have been made, but the way forward is not so clear yet.

Another challenge that needs to be tackled is a decision on the prerequisites for the applicants to the program. We have already taken up the problem of (the lack of) programming and mathematical background. As one of the interviewed teachers stated: "Diversity is an opportunity and challenge".

2.9 Conformance to laws and regulations

The administration keeps up to date with current laws and regulations and integrates these in the routines. It is not clear to the evaluation group to which level and how students and teachers are informed when regulations change.

2.10 Sustainability and equal opportunity

Sustainability was not a requirement when the program was developed. However, some courses do contain examples in courses related to sustainability, e.g., regarding energy consumption. Introduction of sustainability in the curriculum is an issue for further development.

The background of the students is diverse in different ways. The program is open for students from different academic backgrounds. This is seen as very positive as the idea behind the program is to let specialists in other domains learn to use data in their domain to solve problems in their domain.

[3] e.g., `http://dstf.acm.org/`, `https://www.amstat.org/asa/files/pdfs/EDU-DataScienceGuidelines.pdf`

The gender distribution in the program is ca. 40% female and 60% male students. Compared to computer science and engineering programs, this is rather positive. It is, however, noteworthy that there are no Swedish female students. In general, female students are successful although some female students dropped out.

An effort is made to make sure that female students will stay in the program. For instance, there is a mentor program for female students and there is a follow-up on how the female students feel about the program and their studies. Further, the program tries to attract female guest lecturers or tries to collaborate with female specialists for the courses. There is also an effort to attract female teaching assistants that can be role models for the students.

For recruiting purposes the program is advertised using descriptive texts that should be inviting for everyone. There are also movies and interviews taking gender balance into account.

The program is geared towards students with less mathematical background, but many of the optional courses offered by the program require mathematics and programming skills. This is a barrier for many students already enrolled in the program which leads to inequality. This is a challenge for students with less heavy mathematical background. Currently, this is partially solved by allowing students to take courses in programming and mathematics on bachelor level. In practice, this is, however, often not enough as the required bachelor courses are given in Swedish. Some students resolved this by taking courses at other universities. On the one hand we understand this is a problem. Balancing a content that adheres to the different students is challenging because students have different backgrounds, some are more technically oriented than others. Currently, the program attracts many applicants with a wide range of backgrounds, and there is a good balance between male and female students. By increasing the demands on mathematical background there is a risk this balance might not be maintained. On the other hand we believe completion rates might increase and there will be less frustration among students with less technical background. An alternative is to offer a greater range of elective courses whose requirements are inline with the Applied Data Science program. The program leader and teachers already are considering changing the program requirements. The evaluation group recommends to add more courses or sections in courses about the mathematics needed for the program.

2.11 Other observations

The program leadership lacks information that is useful for evaluation and further development of the program. This information relates to, e.g., completion rates for the courses and the program, cancellations of program participation, and how students are doing during their studies. The evaluation group recommends to introduce adequate routines to produce this information to the leadership group.

Although some courses have alumni as guest lecturers or master's thesis supervisors at companies, the alumni are an important source of information

that is not used on the program level. The evaluation group recommends to keep stronger contact with alumni and set up routines for, e.g., alumni meetings and alumni surveys.

There is some inflexibility in the program that seems to be introduced by the way the administration of the program works. For instance, it does not seem possible to change a course after course start (and therefore students choose more courses which results in incorrect information on the number of participants). Further, according to the students and alumni it is not easy to use courses from their bachelor degrees as proofs of fulfillment of prerequisites for courses, which reduces the availability of electives courses a lot.

The program cooperates with a sister program at Chalmers. This has advantages such as more efficient use of personnel and IT systems. There are also disadvantages such as different grading scales, and the fact that some course evaluations are done by Chalmers and some by GU.

3 Summary of the evaluation results

The picture that emerged from the written documents and during the interviews with the program leader, teachers, students, alumni and administrators show that, overall they are satisfied with the form, content and implementation of the program considering that this is a new program. Improvements on different levels are planned. In this section we summarize and highlight some of our observations and repeat the main suggestions for further development of the program.

3.1 Strengths

The program as a whole as well as the courses have good content. The teachers also do a good job in making the content accessible to the students which have very diverse backgrounds.

There is a very good focus on application of data science methods. The program has also managed to create a good network of company collaborations regarding guest lectures, project cooperation and master's thesis projects.

The program manages to provide good education for students from diverse academic backgrounds. Effort is made to bridge the gap between being a person with a previous degree in a field where data science is used to becoming a data scientist (with a possible specialization in the field of the previous degree), thereby opening up new avenues for students. The diversity is also seen as positive and provides an enrichment for the students in the program.

The program leader and teachers are very willing to get comments from students and act upon these.

The gender balance is good and several initiatives are taken to maintain this balance.

3.2 Weaknesses

The fact that the program takes in students from different backgrounds is a challenge, for instance, when it comes to mathematics and programming. For students that do not have a strong technical background there is not much choice regarding elective courses.

Some topics could receive some more attention, such as handling large data, data visualization, soft skills and project management.

There is a lack of data regarding the program and its students that is necessary as a basis for the program leadership to make decisions on further development of courses and the program.

Sustainability is not taken into account except for some examples in courses.

3.3 Future improvements

Applied Data Science is a relatively new program with a huge potential. It is still undergoing many changes and needs to build up a relatively stable curriculum. Inspiration can be found by comparing the curriculum to guidelines for Data Science programs. Contacts with educations programs in other fields where data science is used, may lead to interesting synergy effects. Further, the prerequisites for the applicants to the program need to be discussed based on the experience of the first years of the program and changing some administrative routines may alleviate some problems regarding course selections.

Some of the urgent challenges relate to the need to recruit new staff both for broadening the program content as well as for being able to provide more specialization tracks and enough supervisors and examiners for the master's theses. The popularity of the program may also lead to an increased number of applicants.

To be able to have a good basis for making decisions on the further development of the program, the leadership needs more high quality data. Routines need to be developed to provide the leadership with accurate data regarding, for instance, course participation, course completion and program completion. Furthermore, we recommend to use clear visualizations over the program curriculum with information about the different tracks in the program and the dependencies between the different courses. Also a matrix with courses and program goals is needed to be able to confirm that the program goals are met by all students regardless of which elective courses they choose. The evaluation group also recommends to introduce a reflection report that students should write at the end of their studies which would provide valuable information to the program leadership. Also alumni should be used as an information source. The introduction of a program board with representatives from students and industry is also useful in this respect.

Acknowledgments

The evaluation group would like to thank Birgit Grohe for preparing the material and Lena Kindborg for assisting with practical solutions that helped us carry out the interviews as expected. We would also like to thank the teachers, students, alumni, administrators and student counselor for participating in the interviews.

Appendix A Interview program

8.30-8.45	Introduction meeting with the faculty board and program leader
8.45-9.45	Interview with program leader
9.45-9.55	Pause
9.55-10.55	Interview with teachers
10.55-11.00	Pause
11.00- 12.00	Interview with students
12.00-12.30	Lunch
12.30-13.30	Interview with alumni
13.30-13.35	Pause
13.35-14.35	Interview with study administrators and student counselor
14.35-15.30	Evaluation group summarizes the day and prepares feedback to the faculty and department board
15.30-16.30	Rejoined meeting with the faculty board and program leader
16.30-	Evaluation group work further with the evaluation report

Table 1: Program of the site visit on February 4, 2021

Appendix B Course list

Course name	Compulsory
DIT621 Databases, 7.5 credits	Y*
DIT620 Databases, 7.5 credits	
DIT032 Data Management, 7.5 credits	
DIT856 Applied mathematical thinking, 7.5 credits	Y
DIT855 Applied mathematical thinking, 7.5 credits	
DIT851 Introduction to Data Science, 7.5 credits	Y
DIT852 Introduction to Data Science, 7.5 credits	
DIT861 Statistical Methods for Data Science, 7.5 credits	Y
DIT862 Statistical Methods for Data Science, 7.5 credits	
DIT865 Applied Machine Learning, 7.5 credits	Y
DIT866 Applied Machine Learning, 7.5 credits	
DIT871 Techniques for Large-scale data, 7.5 credits	Y
DIT872 Techniques for Large-scale data, 7.5 credits	
DIT873 Techniques for Large-scale data, 7.5 credits	
DIT875 Research methods for Data Science, 7.5 credits	Y
DIT868 Deep Machine Learning, 7.5 credits	
DIT878 Seminar course in Data Science, 7.5 credits	
DIT891 Project course in Data Science, 7.5 credits	
DIT742 Computational methods in bioinformatics, 7.5 credits	
DIT374 Python for data scientists, 7.5credits	Y
DIT728 Design of AI-systems, 7.5 credits	
DIT245 Machine learning for natural language processing, 7.5 credits	
DIT470 Advanced topics in Machine learning, 7.5 credits	
DIT911 Master Thesis 30 credits	Y
Advanced databases (new)	
Computational methods for large scale data (new)	

Table 2: List of compulsory (marked with Y, Yes) and optional courses. Courses with different course codes indicate the courses underwent a development over time. The Databases course marked with an asterisk is compulsory only for those students that did not take this course, or a corresponding course, during their previous studies.