DD2434 Machine Learning, Advanced Course

Course Analysis, mladv14

Hedvig Kjellström

This analysis has been performed by Hedvig Kjellström in collaboration with Carl Henrik Ek and Jens Lagergren and is based on an online student questionnare with 21 questions, answered by 32 of the students, and on discussions with students during the course. The raw student input is appended to this analysis.

Course Data

The course was given for the first time, during period 2, 2014, with course leader Hedvig Kjellström. 80 Master and 7 PhD students registered for the course, 59 Master and 6 PhD students completed some part of the course, and 49 Master and 2 PhD students are now finished. The teaching activities consisted of:

- 12 lectures of which four were given by Hedvig Kjellström, four by Jens Lagergren, and four by Carl Henrik Ek,
- 5 tutor sessions of which two were given by Jens Lagergren, two by Carl Henrik Ek, and 1 by Hedvig Kjellström,
- Altogether 55 man-hours of oral exam sessions held by Jens Lagergren, Carl Henrik Ek and Hedvig Kjellström, where students presented the home assignments,
- A presentation session where the student presented their projects to each other.

The examination consisted of three home assignments (4 hp) which were performed individually and examined orally in period 2, and one project (3.5 hp) which was performed in groups of four or five and examined by a oral presentation session and a written report in the end of period 2.

The grade distribution on the home assignments is 23 A, 12 B, 8 C, 3 D, and 3 E. The project grades are distributed as 21 A, 16 B, 9 C, 11 D, and 2 E. This leads to a overall course grade distribution of 22 A, 11 B, 11 C, 5 D, and 0 E. PhD students get Pass/Fail.

The feedback from the students is fairly positive. Critical comments are analyzed below, but we would first like to bring forward these comments which warmed our hearts:
”Very good course!!!”
”Overall I think it was one of the best course in my master. I really learned a lot from it.”
”Thank you for this course, I enjoyed it immensely!”
”The things taken up in this course were really good. I would not have wanted to leave KTH without this knowledge.”
”I think this course was really good!”
”Keep the course as a challenge. It was very hard for me, but that steep learning curve was exactly what I was looking for. You responded very well to emails and in the forum, thank you! You put a lot of effort into it and supported the students and where fair, keep that.”

The book used in the course was (K. P. Murphy. *Machine Learning: a Probabilistic Perspective*. MIT Press, 2013). We also used (R. M. Everson and J. E. Fieldsend. A variable metric probabilistic *k*-nearest-neighbour classifier. In *Intelligent Data Engineering and Automated Learning (IDEAL)*, pp 654-659, 2004), (D. J. C. MacKay. Bayesian model comparison and backprop nets. In *Advances in Neural Information Processing Systems (NIPS) 4*, pp 839-846, 1991), as well as a number of papers, one for each project.

What did you think about the book (Murphy)?

1. 12% (5 st) Fantastic!
2. 38% (16 st) OK, I learned a lot from it.
3. 24% (10 st) OK, but would have preferred another book.
4. 19% (8 st) I did not learn much from it.
5. 5% (2 st) It did not give me anything at all.

Main conclusions are listed here, for more info, see below:

- Many students liked that the course was mathematically challenging, while others thought that the course description and prerequisites did not reflect the true level of difficulty. Moreover, the grade distribution indicated that the difficulty level was not too high. **We will maintain the same theoretical level in 2015. However, we will revise the course description and prerequisites to make sure that they reflect the course.**
- The teachers and many students agree that the order in which subjects were presented, and the level of focus on different subjects in the lectures and in the assignments were partially strange. **We will for 2015 revise the order in which subjects are presented in the lectures, and skip some subjects. We will also reduce the number of assignments and rearrange the order of subjects to concorde with the lectures.**
- Most students liked the projects, and the teachers agree that they worked well. Some students felt that some papers were more difficult than others. **We will for 2015 keep the current project format, but introduce more structure to help the students work with the papers, and make informed choices of papers.**
- The opinions about the course book were mixed but mostly positive. **We will keep the book.**

Learning Outcomes

The course is a mathematically rigorous introduction to probabilistic Machine Learning. The learning outcomes and course content were:

Upon completion of the course, the student should be able to

1. explain, derive, and implement a number of models for supervised, unsupervised learning,
2. explain how various models and algorithms relate to one another,
3. describe the strengths and weaknesses of various models and algorithms,
4. select an appropriate model or approach for a new machine learning task.

Machine Learning is the study of algorithms that can improve their performance through experience. Experience usually takes the form of data such as labelled and/or unlabelled examples. Machine learning algorithms are used in a vast number of application domains and tasks. However, to do this successfully a machine learning practitioner must have a systematic understanding of how to learn to perform a required task from data.

It is the goal of this course to give you this understanding. We will present a set of machine learning algorithms and statistical modelling techniques. But more importantly you will learn how the different algorithms are developed, how they are related, and how and when they should be used both in theory and practice.

Most student thought this was clear from the start, and that they had the right prerequisites:

Was the scope of the course clear from the start?

1. 29% (12 st) Yes.
2. 45% (19 st) Relatively clear.
3. 24% (10 st) Not so much.
4. 2% (1 st) Not at all.

Did you have the right prerequisites for the course?

1. 76% (32 st) Yes, completely agree
2. 19% (8 st) Somewhat.
3. 5% (2 st) No, not at all.

The intention has been to give the students a mathematical framework in which a wide range of Machine Learning methods can be formulated, so that they can draw parallels between different methods. We explicitly want to avoid providing a series of cookbook solutions. This puts a lot of focus on math in the course.

From the comments, it is clear that the students who have taken many math courses in their Bachelor had an easier time taking this course. We believe that what makes a difference is the number of math credits the student has taken – in other words, the time the student has spent on studying Math and getting comfortable with mathematical reasoning. Three representative comments:
"For me it was fine but it seems there are students without multivariate calculus that might find it hard."
"I like when it was difficult in the way that you knew what you were supposed to do, just how to do it was difficult. The difficulty "What am I supposed to do?" or "What does htis question really ask?" was sometimes annoying or took too much time."
"Probabilities background was what was missing for me understanding the whole context of the class much easier."

Some students thought there were too many open-ended and non-standardized problems, especially in the project:
"Great experience, but the papers [in the project] has variant difficulties this results in some groups working on easy/hard problem compared to the other groups."
" It took way more time to programm the simulator, than the actual learning algorithm [in the project]."

We would like to make it clear that this is a course in the second cycle (Master level). As opposed to first cycle (Bachelor level) courses, a second cycle course requires that the student takes quite a lot of responsibility. The student should be able to autonomously read scientific books and papers. There will be tasks that are more open than the standardized assignments in the basic computer science and math courses. This is intentional, to train students in structuring and solving complex and diffusely formulated problems.

Other students complain that the lecture notes are not accessible before the lectures, and not readable on their own:
”The lectures where quite fast which is nice but I would like if it was possible to read the slides before attending the lecture.”
”I felt that the slides don't present the information well, especially whole sequences of slides that just had a single figure, with no explanation whatsoever of what is going on.”
”Publish those slides before time, Hedvig's policy on the material releasing time is ridiculous: hand it all out and let the students work earlier if they have time for it.”
We totally disagree. Notes from the lectures is a support for the lecturer during the lecture and for the student’s memory after the lecture. They can not be taken out of context and be read as a free-standing written source of information – if that would have been the intention, the teacher would have written a textbook instead. A message to the students who wrote the comments above: Try reading the parts of the course literature that the teacher always recommends before the lecture!

Action points:

Retain the difficulty level (which we do not think is excessive given the grade distribution – see below) but change the prerequisites to include Multivariate Calculus, add a recommendation to freshen up on the Linear Algebra and Mathematical Statistics basics, and possibly introduce an additional recommendation of a minimum number of Math credits before this course.

Course Activities and Examination

The course activities consisted of 12 lectures and 5 tutor sessions in period 2. The course consisted of three parts, where each part was given by a different teacher and was examined by a separate home assignment. All this corresponded to a course part of 4 hp, graded A-F. This was followed by a project carried out in groups of five students, corresponding to 3.5 hp, graded A-F:

Lecture 1: Introduction Hedvig Kjellström
Murphy Chapter 1

PART 1
Lecture 2: Graphical Models Jens Lagergren
Murphy Chapter 10 (except 10.2.4, 10.2.5, 10.4)
Lecture 3: Mixture Models and the EM Algorithm Jens Lagergren
Murphy Chapter 11 (11.1, 11.2.1-11.2.3, 11.4-11.4.2.5, 11.4.2.8, 11.4.4, 11.4.7), 2.8.1, 2.8.2
Exercise 1: Lectures 2 and 3 Jens Lagergren
Lecture 4: Markov Models and Hidden Markov Models Jens Lagergren
Murphy Chapter 17 (17.3-17.5, 17.6.6)

Lecture 5: Cont. HMM & Undirected Graphical Models Jens Lagergren
Murphy Chapter 19 (19.1-19.3, 19.4.1, 19.4.3)
Exercise 2: Lectures 4 and 5 Jens Lagergren

Oral presentation of Assignment 1

PART 2
Lecture 6: Regression Carl Henrik Ek
Murphy Chapter 2, 5, 14
Exercise 3: Lecture 6 Carl Henrik Ek
Lecture 7: Gaussian Processes Carl Henrik Ek
Murphy Chapter 1, 7, 15
Lecture 8: Learning Representations Carl Henrik Ek
Murphy Chapter 1, 12, 15
Exercise 4: Lectures 7 and 8 Carl Henrik Ek
Lecture 9: Hierarchical Models Carl Henrik Ek
Murphy Chapter 16.5, 28

Oral presentation of Assignment 2

PART 3
Lecture 10: Sampled and Ensemble Models Hedvig Kjellström
Murphy Chapter 16.1-16.4, 23.1-23.5, 24.1, 24.3.7, 1.4.2
Everson and Fieldsend Section 1
Exercise 5: Lecture 10 Hedvig Kjellström
Lecture 11: Topic Models Hedvig Kjellström
Murphy Chapter 2.3.2, 2.5.4, 3.4, 10.4.1, 27.1-27.3

Oral presentation of Assignment 3

Lecture 12: Method and Model Selection Hedvig Kjellström
Murphy Chapter 1, 5.3

Project presentations

We believe that the examination format (no written exam, instead home assignments examined orally) worked well, and we will retain this format for next year. However, we will restructure the assignment series and the order in which subjects are introduced, more below.

We now go through each of the three parts, followed by a discussion about how to restructure the assignment series. After this we discuss the project.

Part 1 (Jens Lagergren)

The idea was that this part would (re)introduce the students to Bayesian probability theory, which would be a good basis for part 2 of the course. However, our assumptions about the mathematical maturity of the students, and the level of knowledge that they acquired in the first ML course (DD2431) were wrong. This meant that the first part was a greater obstacle than we intended, and many students were not able to follow the pace fully.

We asked the students how they found the lectures:

How many lectures did you attend in Part 1 of the course (Lectures 2-5 followed by Assignment 1, held by Jens Lagergren)?

1. 76% (32 st) 3-4.
2. 17% (7 st) 1-2.
3. 2% (1 st) 0.

How did you find the lectures in Part 1 of the course (Lectures 2-5 followed by Assignment 1, held by Jens Lagergren)?

1. 2% (1 st) Excellent.
2. 10% (4 st) Good.
3. 17% (7 st) Ok.
4. 40% (17 st) Not so good.
5. 29% (12 st) A waste of time.

We also asked how they thought the lectures prepared them for Assignment 1:

How well did Assignment 1 align with Lectures 2-5?

1. 40% (17 st) Well.
2. 52% (22 st) Somewhat
3. 5% (2 st) Not very well.

How well did the examination of Assignment 1 reflect what you learned?

1. 33% (14 st) Well.
2. 38% (16 st) Somewhat
3. 26% (11 st) Not very well.

From the answers and comments, we conclude that Assignment 1 was too difficult for many students:
" I was not able to get any nice probability expressions for the Viterbi like algorithm, and it seems like the problem was not worked through before presented as an assignment."
"It was good. The last two problems might have been discussed more thoroughly during lectures. More time was needed for the assignment."
"I think the assignments were good, but way hard for the lower grades. But it was reasonable if you're aiming for A."

The number of students in the course was also much larger than we expected (which we are happy about in general!) which meant that the oral examination became very time consuming for the teachers. The lack of time for each student meant more randomness in the grading. We therefore changed to a more robust grading system, the mean of the best two assignment grades

Moreover, due to a scheduling problem (the schedule is set more than a semester in advance, i.e., long before the course plan was ready) the deadline for Assignment 1 was very close to Assignment 2. This was repeatedly pointed out to the students, but despite this, many students failed to start in time with Assignment 2. More about this below.

Many students comment about the pedagogics of Jens Lagergren. This is not immediately relevant for the course design, but we include some of the comments here to acknowledge that these comments have reached him, and will be taken into account in the design of his future lectures in this course:
"He really explained the math, but I was missing the general idea of the EM-Algorithm framework. It took me a very long time to study and fully understand the concepts from the lecture."

"Did not explain the concepts thoroughly and sometimes seemed confused himself. Should put more explanations on the slides instead of just putting on the word he is talking about without any content. Very good that he had found specific sections in the book for each lecture. It made it feasible to read before the lecture."
"Impossible to understand Jens"
"The lecturer was confused, unprepared, unprepared examples, little or no intuition was given, we wasted time going through derivations."

Part 2 (Carl Henrik Ek)

This was the most successful part of the course, where we introduced a general probabilistic framework to ML. Assignment 2 was very tough for many students, especially given that there were little time – only 3 weeks – between the Assignment 1 and 2 deadlines. This meant that many students got delayed with Assignment 2. The more robust grading system (see above) compensated for that somewhat.

We asked the students how they found the lectures:

How many lectures did you attend in Part 2 of the course (Lectures 6-9 followed by Assignment 2, held by Carl Henrik Ek)?

1. 86% (36 st) 3-4.
2. 7% (3 st) 1-2.
3. 5% (2 st) 0.

How did you find the lectures in Part 2 of the course (Lectures 6-9 followed by Assignment 2, held by Carl Henrik Ek)?

1. 67% (28 st) Excellent.
2. 17% (7 st) Good.
3. 12% (5 st) Ok.
4. 2% (1 st) Not so good.
5. 0% (0 st) A waste of time.

We also asked how they thought the lectures prepared them for Assignment 2:

How well did Assignment 2 align with Lectures 6-9?

1. 67% (28 st) Well.
2. 24% (10 st) Somewhat
3. 5% (2 st) Not very well.

How well did the examination of Assignment 2 reflect what you learned?

1. 60% (25 st) Well.
2. 26% (11 st) Somewhat
3. 10% (4 st) Not very well.

Both the lectures and the assignment were very popular. We will therefore keep this part as is, but move it to the beginning (see below). Three representative comments about the lecturer:
"The lectures were perfect. Intuitive reasonning give more interest to students. Should be the first part of the course. I think it is a good introduction to the course to start with it. I learned more on being baysian at this part than at the first one."

”Really good lectures, you did a very good job. I very much enjoyed them and learned a lot from them. However the assigned readings for each lecture could be more specific. When 2-3 chapters are assigned each time it is simply not feasible to read it all.”
”Carl Henrik is an amazing lecturer!”

Assignment 2 was perceived as hard but good. Three representative comments:
”Really good assignment. Gave a great new look into the theory. I wish I had had time to complete it all, but assignment 1 simply took too much time.”
”I found Assignment 2 harder than Assignment 1 and I also found that the implementation part of Assignment 2 was hard.”
”Very good lab which really helped to understand the content of part 2. Unfortunately I did not have the time to complete the last task of the assignment - maybe a little more time (not competing with assignment 1) would be good.”

Comments by Carl Henrik Ek

This document summarises my reflections of Part 2 of the course DD2434. I will first describe my ideas of what I wanted to achieve with my part of the course and then show how I tried to achieve them. The aim of the second part of the course was to try to show how the elements of machine learning fits together to form a consistent structure. My goal was to give the students a level of abstraction and a framework to allow them to understand and relate different approaches to each other. To get this across I wanted the problems we looked at to be very simple to not ``cloud'' the concepts generality.

Lectures

The three first lectures where focusing strictly on models and the forth where a more abstract lecture introducing hierarchical models. My aim was to separate all the derivations and the intuitions and only talk in very abstract terms during the lectures. The idea was to use very simple data and look at how likelihood, prior, evidence and posteriors interact through Bayes rule. To achieve this a simple one dimensional regression example was used. In the first lecture this was treated as a parametric problem and in the second we used a non-parametric prior. The third lecture introduced latent variable models, in specific we discussed Factor Analysis and introduced the concept of identifiability and therefore the importance of priors to achieve this. The forth lecture discussed neural networks and hierarchical models.

I am happy with the content in the first and second lecture. They were well aligned and there was a continuation building up a story of Bayesian modeling. The third lecture continued this but brought in a few too many things regarding spectral dimensionality reduction which I think should be removed. I think there should be more material relating to identifiability and what it means for inference. This is a good point of talking about inference and also discussing other than Gaussian priors. The forth lecture should be redone completely, it was too abstract for the students. Till next year I will change this and remove the neural network stuff and rather focus on hierarchical priors, a good example would be a mixture model.

Practicals

There were two practical session in this part of the course. During the lectures I tried to avoid derivations and keep the material more abstract. To that end the idea was to collect the essential derivations to these two sessions. During the first part I did go through the derivation of the Gaussian identities. For the second part I had hoped that the students would provide me with something that they needed in order to help them along with the assignment, as this was not the case, I showed them the derivation of a mean-field variational approximation to an Ising model. The motivation behind this was to show them some optimisation and approximations showing how we work with models in practice.

There is need for much more practical exercises for this course for two different reasons. First it would be beneficial for all students to see a few more derivations to get a better feeling for how to work with models. I believe that simply going through all the steps from the lectures but on the blackboard would be useful for many. Secondly, many students lack essential mathematical knowledge and are not at all comfortable working with even trivial mathematical concepts. It is beyond the scope of the course to go through simple linear algebra and elementary concepts of calculus and this is simply something that we cannot provide. However, we can provide pointers of where to look for this and also have open help-sessions where we help the students with what they are working on.

I believe that we also need to filter these short comings of the students down to the basic machine learning course. Concepts such as matrix factorisation, matrix rank, multi-variate Gaussians and basic probability calculus should not have to be covered at all but as the students do not seem to have this knowledge we need to address this in the basic course.

Assessment

The assignment was designed for the students to do at the same time as the lectures. During the lectures there were pointers to the assignment stating how this related to the assignment. After the second lecture the students had the material to pass the assignment and after the third they had the material to get grade C. To achieve grade B or A the students needed to take material beyond what was presented during the lectures. The distribution of grades is

Grade	A	B	C	D	E	F
Number	5	11	5	6	15	15

As grade F are considered students who have passed some assessment during the course but not yet reported assignment 2.

In general I am very happy about this assignment which I feel is also reflected by the course survey. The main comments have been that the assignment were too long and took too much time. However, this assignment corresponded to 1.5 hp which should be one week full time work to pass the assignment. I do think that getting an E should not take more than 10-20 hours which is well within the limits of a weeks work. I also believe that a lot of students had problems simply because they started the assignment too late. The students knew that they had to work on the first two assignments simultaneously but simply did not do this. The lectures followed the structure of the assignments and if they would have worked continuously I do think many would have found it much easier. Till next year this needs to be highlighted and made crystal clear to the students.

There were several comments in the survey regarding students not understanding the questions asked saying that they spent more time to figure out what was asked for than it took to derive the answer. Even though most likely not given as this I consider this as positive criticism. The assignment was aimed at understanding and not at implementing. Therefore actually understanding the question should take most of the effort.

There were some errors in the assignment that will be corrected till next year. I will also add a couple of questions and rephrase certain things to make it more clear. I wanted the students to perform some derivations that I had done during the practicals, for next year I will try to replace them with something easier but something that the students have not seen before. But in general I think that the assignment achieved what we aimed for, learning about three central components of machine learning, priors, posteriors and evidence.

The evaluation of the assignment was done both by a written report and an oral presentation. It was often quite clear what grade the students would get from only the report but being able to explain the assignment and have a discussion with the student was very useful as a learning experience. If it is possible I would like to keep the same structure for next year.

Summary

As a whole I think that the students were happy with this part of the course. Some of them think that the examination was required too much time. I believe that this is due to not having sufficient background and starting the assignment too late. The lecture material should be restructured slightly to keep the same theme rather than branching out to spectral methods in lecture three and neural networks in lecture four. If necessary these things should be done in the basic course. Adding more help sessions and practicals will help students and also providing references to places where they can learn the necessary mathematical background if they find themselves lacking. However, all the mathematical concepts in this course is covered in the basic prerequisite.

Part 3 (Hedvig Kjellström)

The purpose of Part 3 of the course was to introduce different kinds of application, in which the data has different characteristics. This puts different requirements on the methods, and in Assignment 3, the students studied the behavior of parametric methods and non-parametric methods dependent on the data characteristics.

The extent of Assignment 3 was somewhat lowered during the course, when we found that the students spent more time on Assignments 1 and 2 than expected. This meant that not the whole curriculum was examined. This was noted by one student: ”*A lot of the content in the lectures didn't seem relevant for the assignment. A bit confusing.*”.

We asked the students how they found the lectures:

How many lectures did you attend in Part 3 of the course (Lectures 1,10-12 followed by Assignment 3, held by Hedvig Kjellström)?

1. 60% (25 st) 3-4.
2. 31% (13 st) 1-2.
3. 7% (3 st) 0.

How did you find the lectures in Part 3 of the course (Lectures 1,10-12 followed by Assignment 3, held by Hedvig Kjellström)?

1. 29% (12 st) Excellent.
2. 38% (16 st) Good.
3. 21% (9 st) Ok.
4. 7% (3 st) Not so good.
5. 2% (1 st) A waste of time.

We also asked how they thought the lectures prepared them for Assignment 3:

How well did Assignment 3 align with Lecture 10 (Lectures 1,11-12 were not directly related)?

1. 64% (27 st) Well.
2. 31% (13 st) Somewhat
3. 2% (1 st) Not very well.

How well did the examination of Assignment 3 reflect what you learned?

1. 71% (30 st) Well.
2. 24% (10 st) Somewhat
3. 2% (1 st) Not very well.

The students were in general positive to both the lectures and the assignment. Some thought that the assignment was heavy. Others that the third part did not get much attention since many students already had secured their grade from the first two assignments, and only had to do the tasks needed for E in the third assignment:
”With the grading system we ended up with I felt like Hedvig's part got less attention. Now when I think back it feels like I remember more from the first 2/3 of the course, so I do not think you should have a grading system which only contains the grade from two of home assignments.”
”At the lectures I manage to attend I though too much time was spent on describing things that I had seen before and too little time was spent on the things that were new to me. Of course, that is personal. It depends on what you have done before. but I got that feeling from more people in the course.”
”While having pedagogical potential the lectures were confused and repeatedly made deep dives into the trivial parts while glossing ocher the difficult ones.”
”Never attended any but I heard it was good.”
”Hedvig is an amazing lecturer!”
”Nice assignement with just the right amount of work, no "blocker" questions (you could just skip a question and move to the rest and get back to it later), no "read 100 pages in the book to understand one exercise”.”
”I liked assignment3 because it gave me a feeling for the concepts discussed in the lecture. The assignment had the least frustration factor of all of them (I did net get stuck that often).”
”An incredible amount of time went into implementation. The desired answer was not clear from the question.”

Our conclusions for Part 3 is that we should remove subjects and focus on the core concepts. We will integrate Part 1 with Part 3 and let them constitute the second half of the course (see below).

In the assignment, we used AdaBoost as an example of a parametric method. This was a mistake, in that the originally intended focus – what is positive and negative with a parametric vs non-parametric method with different kinds of data? – got lost in implementational details in the AdaBoost implementation. This was also pointed out by some students, e.g. ” *I think this was the best of the 3 assignments, however it had its own issues. The time I spent implementing the weak classifier, and the contour plot if the adaboost decision boundary was longer than the time I spent on part A and B of the assignment. We should be given code for this as to not waste time, since this is really not an important part of the assignment. I liked that the assignment was very heavy on the practical implementation side.*”. Next year we will use a more trivial parametric method, like a perceptron.

Restructuring the assignment series

For next year we will reorganize the order in which subjects are presented, remove some subjects, and lower the number of assignments to two:

PART 1
Lecture 1: Introduction
Lecture 2: Bayesian Linear Regression
Exercise 1: Lectures 1 and 2
Lecture 3: Bayesian Non-Parametrics: Gaussian Processes
Lecture 4: Latent Variable Models: Representation Learning
Exercise 2: Lectures 3 and 4
Lecture 5: Bayesian Inference, Approximate Inference 1
Lecture 6: EM Algorithm
Exercise 3: Lectures 5 and 6

Oral presentation of Assignment 1

PART 2
Lecture 7: Markov and Hidden Markov Models
Lecture 8: Undirected Graphical Models
Exercise 4: Lectures 7 and 8
Lecture 9: Non-Gaussian Models 1
Lecture 10: Non-Gaussian Models 2
Exercise 5: Lectures 9 and 10
Lecture 11: Approximate Inference 2
Lecture 12: Presentation of project papers
Exercise 5: Lecture 11

Oral presentation of Assignment 2

Project presentations

During the first year we wanted to take care of the examination ourselves to increase the chance of consistency in grading. (Examination using several teachers requires quite a lot of additional query design work in order to assure that solutions are interpreted in the same way.) Furthermore, when designing the course, we expected about a third of the number students that we actually got.

For next year, we will engage a couple of PhD students as TAs, and expect the assignments to have matured enough to make examination wit several teachers in parallel possible.

Project

The projects were very successful and many students presented interesting and innovative reports. The students were in general positive to the project:

What did you think about the project (performed in groups of around 5 people, examined with a report and an oral presentation)?

1. 17% (7 st) Fantastic!
2. 71% (30 st) OK, I learned a lot from it.
3. 7% (3 st) OK, but would have been better with a written exam.
4. 2% (1 st) Not so good.
5. 2% (1 st) It did not give me anything at all.

”Great idea to recreate articles. However to me it highlighted that the prerequisites should include more statistics or probability theory since I spend a lot of time explaining basic concepts to my group members.”
”Way to low grades for the project for everyone, otherwise great!”
”The format was great, but all my debugging time could have been put to better use. I don't know if I can say that the method was too hard to implement or what to do about it.”
”I think that the project was good, but it was somewhat difficult to divide the work between all 5 people. I think that a group size of 3-4 people would be more effective. ”

Some students felt that the division into groups based on grade on Assignments 1 and 2 was unfair, while others thought it was good. We think it is fair to let students work with other students at their own ambition level – in that way both the weak and strong students learn the most, and we avoid ”freeriders”. We will retain this principle for next year:
”Groups based on performance might not be the best idea. Just because someone managed to get good grades or was very busy and had little time for the labs does not mean that they will work well together. Even A candidates like to postpone the work until last minute.”
”Remove the meritocracy when creating project groups. Let people choose their groups.”
”The main reason I really enjoyed the project was probably the group I worked with. This is obviously a bit a luck, but I also like the way you made up the groups (based on assignment grades, really nice to be with people on the same level of ambition).”

There were some complaints about group size. While we agree that 3-4 would be easier, the number of students in the course is a prohibitive factor – one teacher can only supervise a limited number of groups.

Some students thought it was hard to select papers, and felt that the paper they got did not live up to their expectations. We will for next year present the papers to the students to enable more informed choices. (Grading works a bit like in Gymnastics: hard problem = more potential for high grade but also more work to finish the project.):
”Great experience, but the papers has variant difficulties this results in some groups working on easy/hard problem compared to the other groups.”
”Ensure that the articles given require a similar effort to replicate. It seemed like this time around, some groups drew a "short straw" and a lot more work than other groups.”

Action points:

Restructure the order of the assingment series, remove some subjects, and lower the number of assignments to two.

Hire TAs.

Present the project papers to the students to help them choose.