Report - DD2421 - 2019-08-30

Respondents: 1
Answer Count: 1
Answer Frequency: 100.00 %

Please note that there is only one respondent to this form: the person that performs the course analysis.

Course analysis carried out by (name, e-mail):

Atsuto Maki (atsuto@kth.se)

COURSE DESIGN

Briefly describe the course design (learning activities, examinations) and any changes that have been implemented since the last course offering.

HT18 (P1) was the third course round to run Machine Learning as DD2421.
The reporting was waited till the end of school year 2018-2019 because the two course rounds, HT18 and VT19, are meant to be basically identical and the students' results can be analysed together.

The course provides an overview of the field of machine learning and describes a number of learning paradigms, algorithms, theoretical results and applications. It also covers some basic concepts of statistics, artificial intelligence and information theory relevant to machine learning.

The course design was basically maintained while there were not urgent need for reforming it.
Lectures, labs, and written exams were held as:

12 lectures with lecture notes made available on course webpage
(6 by Atsuto Maki, 3 by Giampiero Salvi, 2 by Örjan Ekeberg, and a summary lecture by all).
The materials are mainly based on James et al. [1], Prince [2], and Rojas [3] for supplementary reading, all available online. [1] An Introduction to Statistical Learning,
G. James, D. Witten, T. Hastie and R. Tibshirani (Springer, 2013).
[2] Computer Vision: Models, Learning, and Inference, Simon J.D. Prince (Cambridge University Press, 2012).
[3] Neural Networks – a Systematic Introduction, R. Rojas (Springer-Verlag, 1996).

Three lab assignments orally examined by a group of TAs: 1. Decision Trees, 2. Support Vector Machines, and 3. Boosting.

Written exam by A-section (eight multiple choice) + B-section consisting of nine questions. In A-section simple questions regarding basic concept and/or terminology were asked as an essential part for passing. B-Section consisted of questions typically corresponding to a learning outcome (full point is 27). Graded in the range of A-F/Fx.

Changes made from the previous course round (VT18):
• We increased the chances and hours for the drop-in sessions.
• We made a feedback questionnaire on the part of probabilistic machine learning (the results at the bottom of this report*).
• We improved Lab 2 in terms of the instructions, in particular:
- the text has a correction for the error found in the last course round, and
- we have also added a small paragraph with hints on how to debug by using a minimal training dataset
(so that the outcome and intermediate variables can be hand-calculated).

THE STUDENT'S WORKLOAD

Does the students' workload correspond to the expected level (40 hours/1.5 credits)? If there is a significant deviation from the expected, what can be the reason?

The distribution of students' workload corresponds to the expected level, similar to that in P1 in the previous year (compared to P1 rather than P3 assuming that the background of students are similar), while varying between individuals.

THE STUDENTS' RESULTS

How well have the students succeeded on the course? If there are significant differences compared to previous course offerings, what can be the reason?

Overall passing rate was 96% after the re-exam, which was just 1% more than the previous period. Roughly half of the students received 'A' or 'B' and about 15% received 'A' among those, which were the same as before. Overall, there was not significant difference compared to previous course round.

OVERALL IMPRESSION OF THE LEARNING ENVIRONMENT

What is your overall impression of the learning environment in the polar diagrams, for example in terms of the students' experience of meaningfulness, comprehensibility and manageability? If there are significant differences between different groups of students, what can be the reason?

The overall trend in the polar diagrams was similar to VT18, stable as expected. We saw however some improvement (by 1.0 pt) in "15. I could practice and receive feedback without being graded" which was one of the two aspects with a lower rating. This can be regarded as a positive result of intentionally increasing the chances and hours for the drop-in sessions.
Between different groups of students there were no significant difference observed.

ANALYSIS OF THE LEARNING ENVIRONMENT

Can you identify some stronger or weaker areas of the learning environment in the polar diagram - or in the response to each statement - respectively? Do they have an explanation?

Stronger areas stayed strong and a couple of additional areas saw small improvements to have (> 6.0), e.g. "11. Understanding of key concepts had his priority". The weakest area was the same, "20. I had opportunities to choose what to do", but the point has in creased to be above 4.0. The scores themselves, however, may not have much statistical meanings considering the relatively few samples. An improvement was observed in area "15." as explained above.

ANSWERS TO OPEN QUESTIONS

What emerges in the students' answers to the open questions? Is there any good advice to future course participants that you want to pass on?

- It seems to be commonly accepted that the course provides a good introduction to the field of machine learning, which is how it has been developed for students with various backgrounds.
- Not surprisingly, the impressions from a large number of participants are diverse; some would prefer more advanced contents, and others found it hard to follow the formulae. In particular, this was evident with the part of probabilistic machine learning for which we asked for feedback with questionnaire in this course round. The results* (see the bottom of the page) show the challenge to gear the contents towards specific groups of students.

PRIORITY COURSE DEVELOPMENT

What aspects of the course should primarily be developed? How could these aspects be developed in the short or long term?

Through revisions of contents during the seven course rounds since HT14, the course appears to be getting well established. In the medium /long term, however, we continuously take the following aspects into account for further course development:

- Introductions of latest applications of introduced machine learning methods.
- Further clarification of lab instructions if/wherever necessary.
- Re-designing lab 1 to make it more challenging.
- Written exercises in some form.
- Adding another lab.

OTHER INFORMATION
Is there anything else you would like to add?

The TAs were Alexander Kozlov (primary), Sofia Broome, Jesper Karlsson, Taras-Svitozar Kucherenko, Özer Özkahraman, Samuel Murray, Xi Chen, Olga Mikheeva, Tajvar Pouria (doktorand), Tim Lachmann, Thomas Georg Jantos, Evangelia Gogoulou, and Alex Hermansson Grobgeld (MSc students).

* In the questionnaire on the part of probabilistic machine learning we asked "How do you judge?" about the level of the three lectures: Lecture 5 (Probabilistic reasoning), Lecture 6 (Learning as inference), and Lecture 7 (Learning with latent variables). We received 125 responses distributed as below.

----- The lectures -----	Lec 5 / Lec 6 / Lec 7
Way too advanced:	7.3% / 14.4% / 20.0%
Difficult but manageable:	21.6% / 38.4% / 41.6%
Right level for me:	28.0% / 29.6% / 22.4%
Basic but acceptable:	28.0% / 7.2% / 2.4%
Way too basic:	7.3% / 1.6% / .8%
(or, did not attend)	