

Course Analysis for DD2424 - 2022

Josephine Sullivan

1 Course overview

The basic numbers and details for the course:

Administration details

Course Name:	Deep Learning in Data Science
Course Code:	DD2424
Course Credits:	7.5 hp
Distribution of credit:	Homework assignments (4.5 hp), Project (3hp)
Dates of Course:	mid March 2022 - early June 2022

Personnel

Course Leader:	Josephine Sullivan
Additional Teacher:	Hossein Azizpour
Teaching Assistants:	RPL PhD students & undergraduate TAs for homework corrections

Teaching hours within the course

# of lectures:	11 × 2 hour lectures
# of help sessions:	6 × 3 hour sessions

Student numbers and their performance

# of registered students:	293 (but seemed in reality ~250 active students)
# of students completing project:	~240 (after re-exam)
# of students completing assignments:	220
# of students have completed course:	~220 (after re-exam)

2 Goals of the course

The course's objectives are to

- Explain and detail the most common recent architecture networks used in deep learning and how they are applied to different input data.
- Give a coherent and consistent exposition of the back-propagation algorithm so that the students can understand and reproduce how it can be mechanically applied to a wide array of network architectures.
- Introduce to students the common and successful ways to train (supervised + self-supervised approaches) and regularize networks.
- Give the students a sufficiently broad and deep knowledge (and practical experience with DL) to enable student to learn more about the area independently by reading the literature.

3 Changes made to this year's version of the course

Capped the number of students taking the course to 300 This was the fifth year this course has run and was the first year that we capped the number of students to 300 who could take the course. We prioritized students from programmes and/or tracks with DD2424 as a mandatory or conditionally elective course and also required students to be second-cycle students. This meant there was a significant reduction in the number of students taking the course from the previous two years. This was a good thing as it reduced the number of keen but under-prepared students in the course and the workload on me and the TAs. However, there was something strange with many *ghost* students (~ 50) that is students who were listed as being part of the course via Canvas (listed as 293), but then did not complete any activities in the course (233 students completed Assignment 1 and then >210 students completed all 4 assignments which is better rate than in prior years). These ghost students may have been automatic registrations from previous years but not actually active in their KTH studies anymore.

Extended and improved material on self-supervision & Transformers To keep the course up-to-date with recent advances in the field I increased the lecture time devoted to the exciting areas of self-supervision, self-attention and Transformers.

Introduced *Default & Custom* group projects The Group Project was tweaked this year to have two types of projects. One type consisted of “Default Projects” where the skeleton of the project was described to ensure that students successfully implemented well-trained baselines using standard algorithms to get an E grade. Then there were well defined steps to increase their grade from this baseline. The other type of project was the “Custom Project” where the students could define their own project within the scope of the course. The aim was to funnel students with the least experience in software engineering and deep learning into the default projects. With this we could ensure an increased likelihood of proper implementation of standard algorithms for all students (a problem in previous years) and have clearer grading steps. These default projects did give the students and TAs greater clarity and confidence in their results, but they probably need some tweaking to better calibrate the work/level of difficulty needed to increase the grade across the different default projects.

Updated the bonus point exercises for some assignments I increased the difficulty and scope of the bonus point exercises in a couple of the assignments. I think this worked well.

4 Summary of my high level thoughts on the course

I was relatively happy with how the course went. Lectures were conducted in person and I felt the interaction between me and the students during, and before and after lectures was mutually beneficial. As usual though the numbers of students attending the lecture reduced as the course progressed. Not sure if this is unavoidable or whether forcing assessment to cover the lectures later in the course would be beneficial. However, having new assignments entering late in the course would, I think, make students unhappy!

I think most students were pleased with the content of the course and the material presented in the lectures and appreciated the combination of theory and the practical know-how covered in the course. My main aim is to get students to understand the mathematical detail of training neural networks using the back-propagation algorithm and to give them sufficient knowledge of deep learning to read and understand research papers in the field. And I think this is achieved as the students who pass DD2424 and go onto take the Advanced Deep Learning course DD2412 seem to cope well.

I think the majority of students would prefer the programming assignments to be better supported for python users. The students can complete the assignments

in python and/or Matlab but the instructions are written assuming the students are completing them in Matlab.

5 Teaching

The teaching in the course consisted of traditional lectures in tandem with help sessions manned by the TAs.

6 Assessment in the course

To pass the course the students had to

- complete the programming assignments and upload a report on each assignment **and**
- complete a group project, write a report on it and make an oral presentation of their project

The students' programming assignments were reviewed by the TAs and myself and were graded as pass/fail. The projects reports were graded from A-F and this grading was performed by me.

7 Reading list

The original textbook for the course [Deep Learning](#) by Ian Goodfellow, Yoshua Bengio and Aaron Courville has become a little dated. The textbook was used more of a complement to the course as the much of the material is self-contained. Also it is freely available online. The book still obviously contains useful material but the field is moving so we also point the students to the book *Dive into Deep Learning* by Aston Zhang, Zack C. Lipton, Mu Li, and Alex J. Smola.

8 Requirements

There are no formal pre-requisites, beyond that the students have taken 45 credits in mathematics or informatics, for the course but we should explicitly recommend the students to have taken the course: *Machine Learning* DD2431 beforehand as well as courses in programming and linear algebra.

9 Actual course content

The schedule is available at the [KTH Canvas webpage for DD2424 2022](#) and shows the exact topics covered by the course.

10 Planned changes for next year

10.1 Aspects of the course that should be improved

Investigate using Julia as the programming language for the assignments It would be nice to have one programming language for the assignments. Matlab and python both have a list of pros and cons as being this single language. One potential option as it potentially has quicker run time than python/Matlab and I think has easy access to auto-differentiation methods. The latter would allow students to more easily debug their code. The plan is to investigate the feasibility of this option.

Replace/add an assignment based on self-supervised learning

Assignment 3 of the course is a little bit fiddly and not hugely satisfying for the students. As self-supervised training is becoming such an integral part of deep learning, it would be very cool to develop an assignment based on this concept. It would be fun to build a non-trivial but computationally feasible assignment based on contrastive learning. This idea will be explored with the aim of perhaps introducing an option to complete this new assignment instead of the existing assignment 3.

Improve the specification of the default projects and the grading steps

The introduction of the default projects worked well and in coming year we will develop this idea and tweak the grading steps so there is better calibration of the grading criteria between the different default projects.

GPU resources??? Once again some students encountered problems accessing GPUs on via GCP. Google was not at all concerned with helping us get to the bottom of the problem! At the moment there is not a clear resolution to this problem with GCP or obvious options for other GPU resources. Fingers crossed that EECS/KTH can come to the rescue some how.....

10.2 Summary of planned improvements

Here is a list of the planned improvements for the next year's course:

- Investigate if using Julia instead of Matlab/python for the programming assignments would be a good idea.
- Better calibration and perhaps a little more diversity in the *Default Projects*.
- Hope that EECS/KTH develops some large scale infrastructure for GPU usage for second-cycle students.
- Look into creating an assignment based on self-supervised training.