Course Analysis for DD2424 - 2023

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 2023 - early June 2023

Personnel

Course Leader:	Josephine Sullivan
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:	236 (but seemed in reality ∼195 active students including P
# of students completing project:	∼160 (after re-exam)
# of 2nd cycle students completing assignments:	163
# of 2nd cycle students have completed course:	∼160 (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

Extended and improved material on self-supervision, LLMs & Diffusion models   To keep the course up-to-date with recent advances in the field I increased the lecture time devoted to the exciting areas self-supervision and large-language models. I also created a presentation describing diffusion models for image generation.

Updated the bonus point exercises for some assignments   Once again tinkered with the bonus point exercises in the assignments to keep them interesting and up to date. The keen students appear to enjoy these exercises.

Looked into changing the progamming language to Julia   I abandoned the idea of converting to Julia as it become clear that it is probably still too early to use Julia for deep learning. The libraries for performing deep learning are manifold have not stabilized and have some issues. Perhaps this idea can be revisited in a couple of years.

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. The book 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 now formal pre-requisites for taking DD2424. These correspond to having passed courses in the following:

- Programming
- Linear algebra
- Single and multi-variate calculus
- Probability and statistics
- Machine Learning OR Artificial Intelligence

9   Actual course content

The schedule is available at the KTH Canvas webpage for DD2424 2023 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

Replace/update assignment 3

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 have been very cool to develop an assignment based on this concept. However, it was not feasible to build a non-trivial and computationally assignment and still involve the cifar-10 dataset. Thus I will investigate other potential ideas this year. Foremost will be to either to replace batch-normalization with layer normalization as or be more ambitious and investigate some form of supervised contrastive learning.

Improve once again the specification of the default projects and the grading steps

The introduction of the default projects has worked well. Once again though I would like to tweak the grading steps so there is better calibration of the grading criteria between the different default projects and it is a bit more challenging to get the different grades. Once again it is always a trade-off to be made

GPU resources

There were fewer issues with GCP this year. It is unclear whether this was because fewer students took the course, GCP was easier to use or more students relied on their own hardware. I will keep an eye on this.

10.2   Summary of planned improvements

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

- Provide python versions of the matlab scripts already provided.

- For the *Default Projects* continue with improved calibration between these projects and add slightly more difficulty for each of the grade steps.

- Update the list of potential *Custom Projects*.

- Make another attempt to replace/change assignment 3.

- Update relevant lectures to keep up with the developments in AI that have occurred since the last version of the course.