

Practical Session 1: Hyper-parameters and training basics with PyTorch

Cyriaque Rousselot, Matthieu Nastorg,
Francesco Pezzicoli, Guillaume Charpiat *

24th January 2023

This first Practical Session aims at giving you some understanding about the various hyper-parameters that need to be tuned when training a deep learning model. This implies the conception of the model architecture as well as the specifics of the training process. Each problem has its particularities, and needs a tailored model to solve it.

The exercise is split into 2 problems, each presented as a Jupyter notebook. For each of them, your task will be to gradually experiment with the various accessible hyper-parameters to create a model that can solve this problem with good accuracy and reasonable computing resources.

You can do this work in groups composed of up to 3 persons. Each group will have to submit a written report (.pdf) of 2 pages containing the following:

- For each problem, a paragraph including the architecture details of the best model you managed to conceive, and the training process used to train it, as well as its numerical performance;
- A global discussion (for all problems together) including:
 - Your analysis of the trade-offs between the performance of the final model, the time it takes to train it, and the time needed to tune all the hyper-parameters;
 - For each relevant kind of parameters (the various architectural choices, the optimiser configuration, the training process details, etc..), a paragraph explaining what you found about how it impacts the training process and the performance of the model.

*based on materials conceived with Victor Berger & Wenzhuo Liu

This homework is to be sent within 2 weeks, i.e. before February 7th. Do not wait for the last moment to start, as hyper-parameter tuning is time-demanding. Playing with hyper-parameters might seem a bit tiresome but it is a necessary step to gain experience with them, understand them and be at ease later for the training of bigger networks.

More information:

- The report (in pdf format) will have to be submitted on a website that will be announced on the course webpage as well as on the Discord forum (do not forget to join it!). Please aim at 2 pages as asked (not 1 nor 5!).
- A forum is opened (on Discord), for students to help each other on technical points and to ask questions to teachers and TAs (link received when registering to the course).
- Possible environments to consider are proposed on the course webpage (local installation vs. remote free computing server) . Note that all exercises will be conceived so that your laptop's CPU is sufficient (no GPU card needed).
- On the course webpage you will also find a document entitled Tips and tricks to train neural networks, containing general advise for neural network training. It might help to read it again each time you encounter an issue.