

Practical Session 1: Hyper-parameters and training basics

Victor Berger Wenzhuo Liu Guillaume Charpiat

Monday, January 13th, 2020

This first exercise 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 4 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 of 3-4 pages containing the following:

- For each problem, 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
- 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

This homework is to be sent before next session, which takes place on February 3rd. This means you have 3 weeks to accomplish it, but do not wait for the last moment to start, as it was designed to last 3 weeks!, it will be 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. A website to submit your report will be announced on the mailing list (do not forget to subscribe to it!). A forum is also opened, for students to help each other on technical points and to ask questions (link on the course webpage, password: mva-dlip).

For students already very familiar with these hyper-parameters, an optional, supplementary exercise targeting other hyper-parameters will be also proposed (NB: the 4 first problems will still need to be solved, this is just a bonus).

Links:

- First problem (the 3 other problems will be announced shortly):
https://www.lri.fr/~gcharpia/deeppractice/2020/TP1/Training_binary_classification_model_with_pytorch.ipynb
- Tips and tricks to train neural networks:
<https://www.lri.fr/~gcharpia/deeppractice/2020/tips.pdf>