Homework 1: Supervised Deep Learning

General overview: In this homework you will learn how to implement and test simple neural network models for solving supervised problems. It is divided in two tasks. The *regression task* will consist in a simple function approximation problem, similar to the one discussed during the Lab practices. The *classification task* will consist in a simple image recognition problem, where the goal is to correctly classify images of Zalando's article images (FashionMNIST).

In both cases, you should explore the use of advanced optimizers and regularization methods (e.g., initialization scheme, momentum, ADAM, early stopping, L2, L1 / sparsity, dropout…) to improve convergence of stochastic gradient descent and promote generalization. Learning hyperparameters should be tuned using appropriate search procedures, and final accuracy should be evaluated using a cross-validation setup. For the image classification task, you can also implement more advanced convolutional architectures and explore advanced feature visualization techniques (e.g., visualization of feature maps and/or input optimization by maximizing the activation gradient) to better understand how the deep network is encoding information at different processing layers.

Additional information and details related to the datasets are provided in a Python notebook uploaded in the course Moodle.

Technical notes: The homework should be implemented in Python using the PyTorch framework. The student can explore additional libraries and tools to implement the models; however, please make sure you understand the code you are writing because during the exam you might receive specific questions related to your implementation. The entire source code required to run the homework must be uploaded as a compressed archive in a Moodle section dedicated to the homework. If your code will be entirely included in a single Python notebook, just upload the notebook file.

Final report: Along with the source code, you must separately upload a PDF file containing a brief report of your homework. The report should include a brief Introduction on which you explain the homework goals and the main implementation strategies you choose, a brief Method section where you describe your model architectures and hyperparameters, and a Result section where you present the simulation results. Total length must not exceed 6 pages, though you can include additional tables and figures in a final Appendix (optional). The report must be self-contained, so please make sure that all relevant results are explicitly included in the report / notebook.

Grade: The maximum grade for this homework will be **8 points**. Points will be assigned based on the correct implementation of the following items:

- 2 pt: implement basic regression and classification tasks
- 2 pt: explore advanced optimizers and regularization methods (both tasks)
- 1 pt: optimize hyperparameters using grid/random search and cross-validation (both tasks)
- 2 pt: implement CNN for classification task
- 1 pt: visualize weight histograms, activation profiles and receptive fields

Deadline: The complete homework (notebook or source code + report) must be submitted through Moodle at least 10 days before the chosen exam date.