Teachers: J. Hennebert, M. Melchior
C. Gisler, P. Matvulj, O. Zayene

Practical Work 07 – 04/04/2024
Back-Propagation

Objectives

Main objective is diving into the mechanics of the back propagation algorithm and implement it for an MLP. Again, the solutions can be elaborated by starting from a Jupyter notebook that will give you guidance.

Another goal is that you learn how GradCAM works and that you implement and apply it to MNIST.

In this PW, we require you to provide implementations in pytorch.

Submission

— **Deadline** : Wednesday 17 April, noon

— **Format** :

Completed and well documented Jupyter notebook with the blanks filled in.

Indicate the name of the group in the name of the notebook.

— **One submission per group !**

Exercise 1 Implementation of Back-Propagation without Autograd

Implement the back-propagation algorithm for an arbitrary MLP by using pytorch tensors but **without using pytorch's autograd functionality, i.e. no** *backward*()**-call on a pytorch tensor**.
Use as a basis the Jupyter notebook `pw07_backprop_stud.ipynb`. The notebook should guide you through all the steps needed.

Finally, you will be able to

a) validate whether you compute correct gradients by comparing with the results provided by pytorch's autograd ;

b) train an MLP with an arbitrary number of layers. You should achieve similar performance as obtained in PW03.

Exercise 2 Optional : Autograd

Watch the two videos sessions 4.1 and 4.2 from Francois Fleuret's deep learning lectures (https ://fleuret.org/dlc/). Play through at least some of the examples.

Exercise 3 GradCAM

GradCAM is a popular tool to visualise where a model has been looking at when making predictions. It is one of the tools in a typical XAI toolbox.

Study the video and slides of session 9.3 from Francois Fleuret's deep learning course (slides here : https ://fleuret.org/dlc/materials/dlc-slides-9-3-visualizing-in-input.pdf) where you will find details on how GradCAM works and how it can be implemented.

Implement GradCAM for a model of your choice (minimum 2 CNN Layers, 1 dense layer) for Fashion-MNIST classification. Use the last layer of your CNN-stack to define the GradCAM.

Analyse the results for some selected images (for each class at least one sample). Where is GradCAM providing explainability - where not ? Can you identify regions highlighted by the GradCAM that are specific to the digits ?

Prepare a notebook with the implementation and the analysis for submission.

Exercise 4 Optional : Review Questions

a) Explain the difference between gradient descent and back-propagation. What are the benefits of using the backprop algorithm for computing gradients ?

b) What are the variables in MLP that need to be remembered per layer during the forward pass to have them available in the backward pass ?