

Neural Networks from Scratch
Winter Term 2024

Exercise 5

- Return electronically until Tuesday, December 17, 2024, 09:00
- Include your names on the top sheet. Hand in only one PDF.
- A maximum of three students are allowed to work jointly in a group.
- When you include plots add a short explanation of what you expected and observed.
- Hand in source code if the exercise required programming. You can bundle the source code along with the PDF in a .zip file.

5.1 Project Work: Transformer from Scratch

Complete the nanoGPT implementation, which can be found in the template folder. For this you will likely need to install the *icecream* and *tokenizers* packages. These can be installed using *pip* or *conda*. In particular implement the backward and update routines of the following components:

- GoePT (nanoGPT class)
- Dropout
- LayerNorm
- GELU
- MLP
- MultiHeadAttention
- Embedding
- Block
- Cross-Entropy Loss (not explicitly a backward component)

These are located in the following files of the template: *models/GoePT/loss.py* , *models/GoePT/model.py* and *models/GoePT/layers.py* *model.py* contains the main function of the code. The template should be run from the root of its folder, as example with: *python models/GoePT/model.py*

To verify that your implementation works, run it for at least five epochs and verify that the loss decreases. Plot both the training and validation loss over the number of epochs. The whole implementation should be done either in *NumPy* or *CuPy*. We recommend using *NumPy* for the initial development.

Note: You may also find writing GPU agnostic code in *CuPy* interesting for this project.¹

If you find any bugs in the template, please inform us as soon as possible, as this likely affects all groups.

5.2 Individual Group Project Proposal

For the second part of the project work, please propose a project, which you'd like to tackle as a group. As inspiration you can take a look at the examples presented in the lecture. The proposals will be discussed on December 18th to make sure they are in scope for a practical.

Your proposal should contain the following:

- Title of the project
- Abstract explaining the scope, goal and work to be done in the project.
- Estimated timeplan for the project.
- Estimated work distribution within the team.

¹https://docs.cupy.dev/en/stable/user_guide/basic.html#how-to-write-cpu-gpu-agnostic-code