Heidelberg University
Institute of Computer Engineering
Computing Systems Group

Franz Kevin Stehle
Hendrik Borras

Neural Networks from Scratch
Winter Term 2024

Exercise 4

- **Return electronically until Tuesday, November 19, 2024, 09:00**

- **Include your names on the top sheet. Hand in only <u>one</u> 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.**

4.1 Reading

Read the following paper and provide a review as explained in the first lecture (see slides):

- Ashish Vaswani, et al. 2017. Attention is all you need.

Choose one of the following two papers and provide a review as explained in the first lecture (see slides):

- Alexey Dosovitskiy, et al. 2020. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.

- Mehdi Ali, et al. 2023. Tokenizer Choice For LLM Training: Negligible or Crucial?