Problem Set 12, Dec. 2, 2024 (GPT)

Goals. The goal of this exercise is to

- Familiarize yourself with the GPT models.

- Have time to discuss Project 2 with the assistants and teammates.

1   Train GPT to perform multiplications

The goal of this exercise is for you to become more familiar with transformers and GPT. We will focus on a much simpler task than language modeling that can be trained in a few minutes. Specifically, we will train a small GPT model from scratch to perform multiplications. We will use individual characters as tokens. Doing multiplications directly for example in the form: "12*34=408" can be challenging. Even large language models are often not able to do this accurately for moderately large numbers, say 5 digits (although it is possible using special tricks). Use the following Jupyter notebook gpt-multiplication.ipynb.

Open in Colab: colab.gpt-multiplication.ipynb. This gives you access to a free GPU.