# **Chapter 5: Pretraining on Unlabeled Data**

### **Main Chapter Code**

* [ch05.ipynb](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/ch05.ipynb) contains all the code as it appears in the chapter
* [previous\_chapters.py](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/previous_chapters.py) is a Python module that contains the MultiHeadAttention module and GPTModel class from the previous chapters, which we import in [ch05.ipynb](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/ch05.ipynb) to pretrain the GPT model
* [gpt\_download.py](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/gpt_download.py) contains the utility functions for downloading the pretrained GPT model weights
* [exercise-solutions.ipynb](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/exercise-solutions.ipynb) contains the exercise solutions for this chapter

### **Optional Code**

* [gpt\_train.py](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/gpt_train.py) is a standalone Python script file with the code that we implemented in [ch05.ipynb](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/ch05.ipynb) to train the GPT model (you can think of it as a code file summarizing this chapter)
* [gpt\_generate.py](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/gpt_generate.py) is a standalone Python script file with the code that we implemented in [ch05.ipynb](https://github.com/rasbt/LLMs-from-scratch/blob/main/ch05/01_main-chapter-code/ch05.ipynb) to load and use the pretrained model weights from OpenAI