

Module 7

Programming Assignment

Programming assignments are due by 11:59:59 PM on the day of the next lecture. All programming assignments are graded for clarity and functionality. Your code should be well-organized and include helpful comments where appropriate. Please upload your completed assignment to canvas as a zip file.

In this assignment, you will implement a sampling object which is passed logits from a GPT-style model and selects a token to generate. If used repeatedly, this can be used to generate text.

Your objective is to complete the Sampler class in **sampler.py**, which should include implementations of Top-k and Top-P sampling (mutually exclusive), and both a Frequency penalty and a Presence penalty (not exclusive).

For a practical example of using Sampler, see the additional file **example_use_with_gpt2.py**

Your deliverables for this assignment are:

- A completed *sampler.py* containing the definition of the Sampler class.