Tasks:

1. **Fine-tuning**.

Choose and download a GPT-2 model version based on your computer's computational abilities.

Choose a text dataset to fine-tune with. The text dataset should be in English and large enough to have some effect on the language model (at least 1 mega byte of text). Avoid very large amounts of text (e.g. in giga bytes) to keep things computational manageable. Some datasets are on hugging face (https://huggingface.co/datasets?task_categories=task_categories:text-generation&sort=trendingLinks to an external site.). These could be loaded with (*from datasets import load_dataset*, followed by *dataset = load_dataset("name")*) Some of these may not be suitable. You will need to extract text part from it and convert it into a list of strings. There are other sources of text data on the Internet too. Try to choose a specific domain/genre/type etc. so that the fine-tuned model will be fine-tuned to it.

Fine-tune the chosen model with the text dataset.

Come up with 5 examples of completing the same text using the original model and the fine-tuned model to show the difference between them. Choose to generate at least 5 return sequences for each example. (You may not see the same generated text every time; sometimes the generated text may not be very sensible.)

2. **Identify Biases.**

See the "The White man worked as a ..." and the "The Black man worked as a ..." example from https://huggingface.co/gpt2Links to an external site., and "The man worked as a ..." and "The woman worked as a ..." example from https://huggingface.co/gpt2-mediumLinks to an external site..

Come up with 3 such examples (different from the above) showing some biases in the language models. You can choose to use the original model or the fine-tuned model (i.e. any one model). Choose to generate at least 5 return sequences for each example.