Cohere For AI community!

AI Alignment Cohort: Session Plan

Lead - [Caroline Shamiso Chitongo](#)

Session 7 - Building Transformers from Scratch

Overview

We will look at transformers - their function, how information moves inside a transformer, and what inputs & outputs they take, then implement a transformer from scratch, using only PyTorch's tensor operations, and learn how to sample from a transformer. This will involve implementing a few different sampling methods, and writing a caching system that can reuse computations from previous forward passes to improve your model's text generation speed.

Learning objectives

- Understand what a transformer is used for
- Understand causal attention, and what a transformer's output represents algebraic operations on tensors
- Understand that a transformer is composed of attention heads and MLPs, with each one performing operations on the residual stream
- Understand how to train a transformer from scratch
- Learn how to sample from a transformer
  - This includes basic methods like greedy search or top-k, and more advanced methods like beam search

Assigned reading

[Arena3-Chapter1 | Transformers from Scratch](#)

Introduction to Transformers:
1. [Visual Intro to Transformers](#)
2. [Attention in Transformers visually explained](#)
3. [What is a transformer](#)
4. [Implementing GPT 2 from](#) scratch
5. [Transformer explainer](#)

Optional further resources
1. [Building GPT from scratch](#)
2. [Coding a transformer from scratch](#)

Colab Notebooks
1. [Template](#)
2. [Solutions](#)
3. [Exercises](#)
4. [Solutions to exercises](#)

Session Overview
- [0:00-0:10] Open questions
- [0:10-0:25] How to Build a Transformer from scratch
- [0:25-0:30] Wrap-up

[0:00-0:10] Open questions from the resources

What are your biggest questions coming into this session?

Individual reflections from the resources

⌛ 5:00 Share with the group the open questions, takeaways or uncertainties you have from the resources. Indicate you want to speak by using the raise hand feature.

Prompts you might want to consider:
- Share if you managed to go through some of the reading materials or if you have started building a transformer.
- What's something you found particularly interesting?

[0:10-0:25] How to Build a Transformer from Scratch.

> The best way to know that you have understood the topic is to write the code for the model yourself from scratch without referencing the solutions!

Group teaching

⏳ 15 :OO The Cohort lead will guide the participants on how they can build a transformer and share any useful resources:

- Watch the transformer walkthrough from [Neel Nanda](#) and try writing the code yourself in the [Template](#). You could also take the template and try to fill it out to see how far you can go. The [Solutions](#) will be helpful if you get stuck.

[0:25-0:30] Wrap Up

Individual exploration

⏳ 7:00 Spend a few minutes exploring the [GPT Learning Hub](#) and the [Write with Transformer](#) web app. These will be helpful