

Building a GPT in Stan

StanCon 2023



mc-stan.org

Instructor: Daniel Lee
bearlee@alum.mit.edu

Course description

GPTs have captivated the world. If you want to understand what it is, we'll be programming one from scratch. We'll look at the Generative Pre-trained Transformer (GPT) from a Stan perspective and understand the basics by building the statistical model in stages. This will be a good use of time if: a) you've heard of GPTs, but haven't dug into the literature, b) you've read the literature but haven't had the time to implement a GPT, or c) you're a Stan user and want to get better through a hands-on workshop. The only prerequisite is that you know some Stan programming. (We'll be loosely follow Andrej Karparthy's Let's Build GPT: from scratch, in code, spelled out. <https://youtu.be/kCc8FmEb1nY>.)

Course length: 1/2 day