

Tutorial at StanCon 2023

Building a GPT in Stan

Description

It took me a while to link the concepts in literature to the underlying statistical model of the transformer. My hope is to shorten that journey for you. Over the half-day tutorial, we'll be building up a Generative Pre-trained Transformer (GPT) model from scratch. We will be implementing the model in Stan while cross referencing the literature. The hope is that we can all leave with an idea of what the statistical model that corresponds to GPT is.

What this tutorial will not cover

- A production level GPT; Stan wasn't designed for this purpose.
- A deep dive into large language models. This is a fast moving field and by the time this tutorial happens, the field will have advanced.

What this tutorial is

- A space to work on your Stan skills. 3 hours building 6 different Stan programs. It's going to be work to keep up. You'll get a chance to learn different tips and tricks to debugging programs quickly.
- A guided tour of the concepts in the LLM literature and how we can think about it as we implement concrete Stan programs. Hopefully this tutorial will demystify the GPT model and translate the terminology to things that are recognizable. It may even inspire you to bring some of these techniques back to your applied work.
- Informal and fun. I mean... if you think of this as fun, please come and join us!

Prerequisites (suggested)

- Please know some Stan programming.
- A willingness to try. If you want to sit in and have it be like a (slow) cooking demonstration, that's perfectly fine. There will be code available at each checkpoint.

Instructor

Daniel Lee is a computational Bayesian statistician who helped create and develop Stan. He has 20 years of experience in numeric computation and software; over 10 years of experience working with Stan; and has spent the last 5 years working on pharma-related models including joint models for estimating oncology treatment efficacy and PK/PD models. Past projects have covered estimating vote share for state and national elections; satellite control software for television and government; retail price sensitivity; data fusion for U.S. Navy applications; sabermetrics for an MLB team; and assessing "clutch" moments in NFL footage. Daniel has led workshops and given talks in applied statistics and Stan at Columbia University, MIT, Penn State, UC Irvine, UCLA, University of Washington, Vanderbilt University, Amazon, Climate Corp, Swiss Statistical Society, IBM AI Systems Day, R/Pharma, StanCon, PAGANZ, ISBA, PROBPROG, and NeurIPS. He holds a B.S. in Mathematics with Computer Science from MIT, and a Master of Advanced Studies in Statistics from Cambridge University.

Please email any questions to: daniel@bayesianops.com