

GPT Mastery: A Comprehensive Guide To Crafting Your Own ChatGPT

Week 7

Hey everyone! This is the final video of this course, and it's a very important one where we explore how GPT-2 works.

1.) Let's reproduce GPT-2 (124M)

In this video, we recreate the GPT-2 model (124M) from scratch, starting by building the GPT-2 network, then optimizing its training for speed, and setting up the training run following the GPT-2 and GPT-3 papers' hyperparameters. Once everything is set, we initiate the training and check back the next day to review the results and enjoy some amusing model generations. This video builds on knowledge from earlier videos in the Zero to Hero Playlist on my channel, and by the end, it resembles my nanoGPT repo by about 90%.

Youtube:  [Let's reproduce GPT-2 \(124M\)](#)

Github: <https://github.com/karpathy/nn-zero-to-hero>