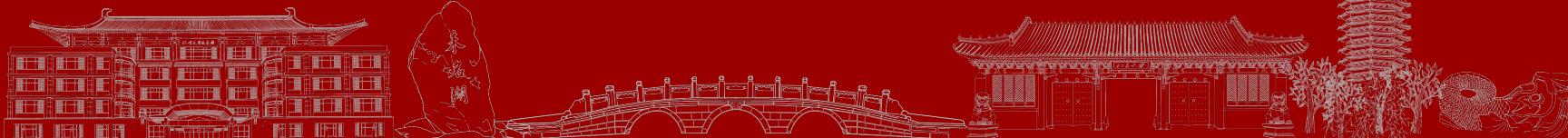




训练模型、训练人生

尹训健

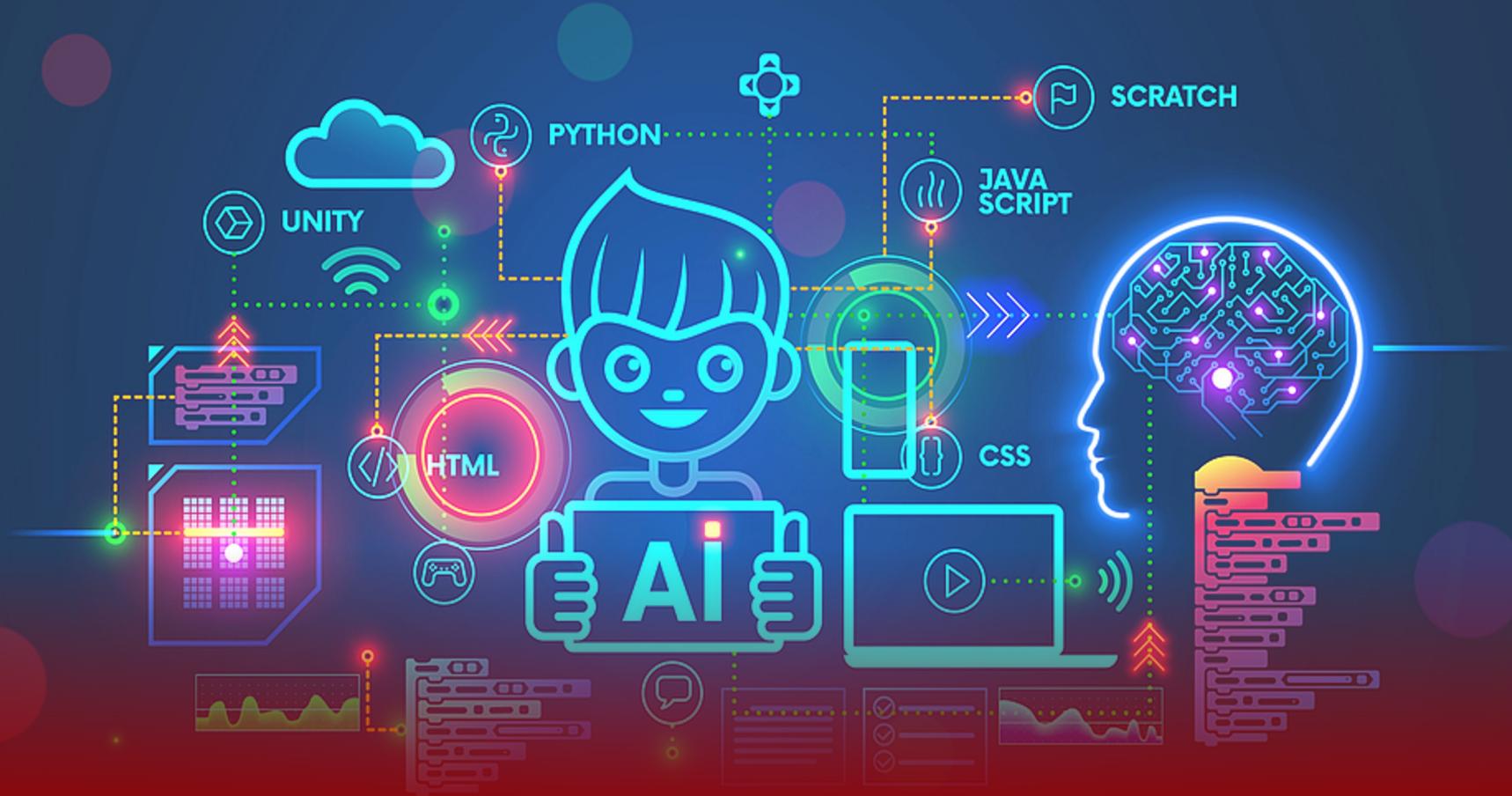
前沿交叉学科研究院2022级硕士研究生



The background features a man in a dark suit and a blue hard hat standing on a dark, textured surface, possibly asphalt, under a dark sky. He is looking upwards towards the horizon. The overall atmosphere is mysterious and futuristic.

ChatGPT

ChatGPT的出现，彻底改变了各行各业，
也改变了我们。



我们总在研究怎么让AI更像人，
我们能不能反过来思考，向AI学点东西

Next Token Prediction

&

Scaling Law

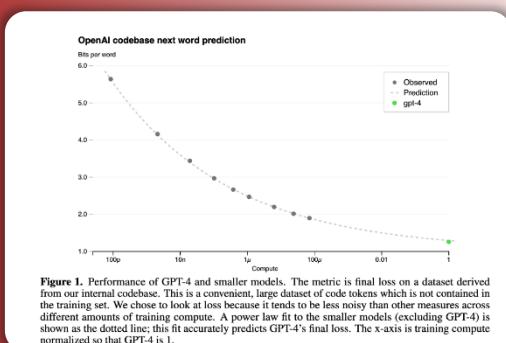
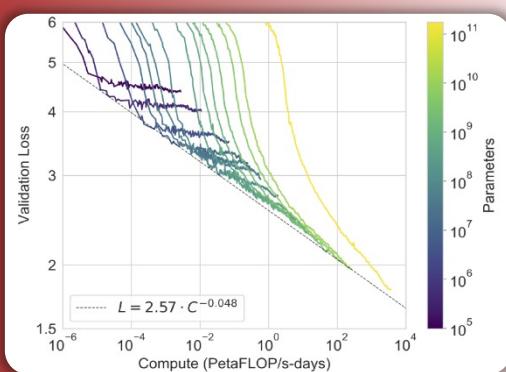


Figure 1. Performance of GPT-4 and smaller models. The metric is final loss on a dataset derived from our internal codebase. This is a convenient, large dataset of code tokens which is not contained in the training set. We chose to look at loss because it tends to be less noisy than other measures across different amounts of training compute. A power law fit to the smaller models (excluding GPT-4) is shown as the dotted line; this fit accurately predicts GPT-4's final loss. The x-axis is training compute normalized so that GPT-4 is 1.



“你是一
个极为优秀
的数学专家”



“生命以痛吻我， 我却报之以歌”

说了这么多次，怎么还是理解不了我的意思！

抱歉让你觉得我没get到你的意思！ 😭 能不能再具体点说说你想表达啥，或者换个方式描述一下？我会努力跟上你的节奏，给你最贴心的回答！ 😊

⟳ ⏺ ⏹ ⏻ ⏸ ... 2.6s



谢谢北大，谢谢荷院。

我们的预训练阶段已经结束，
但这也是持续学习与迁移学习阶段的开始。

