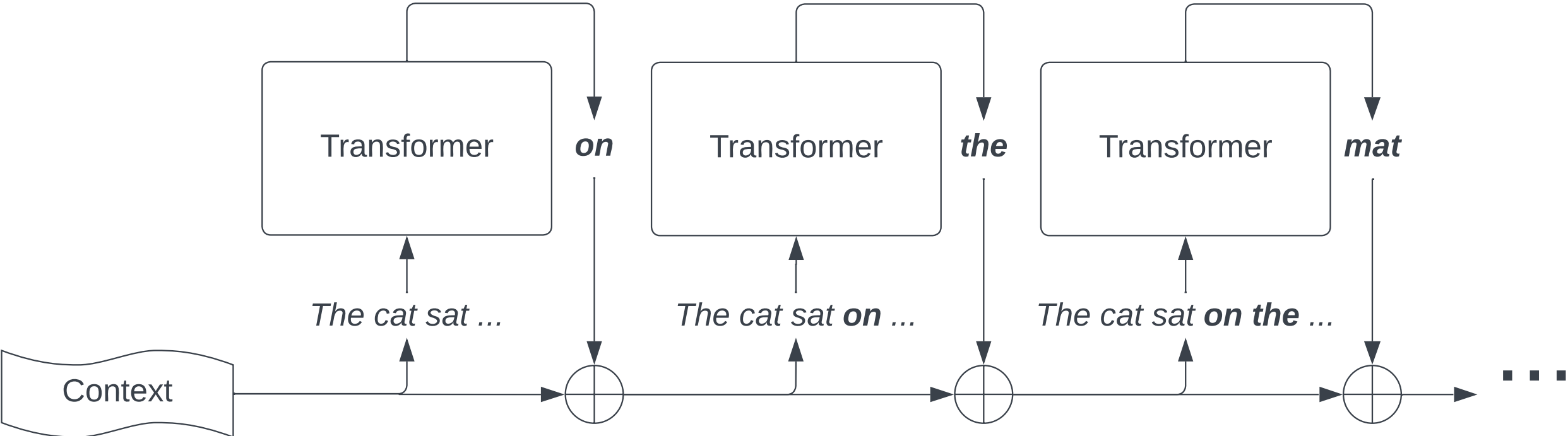
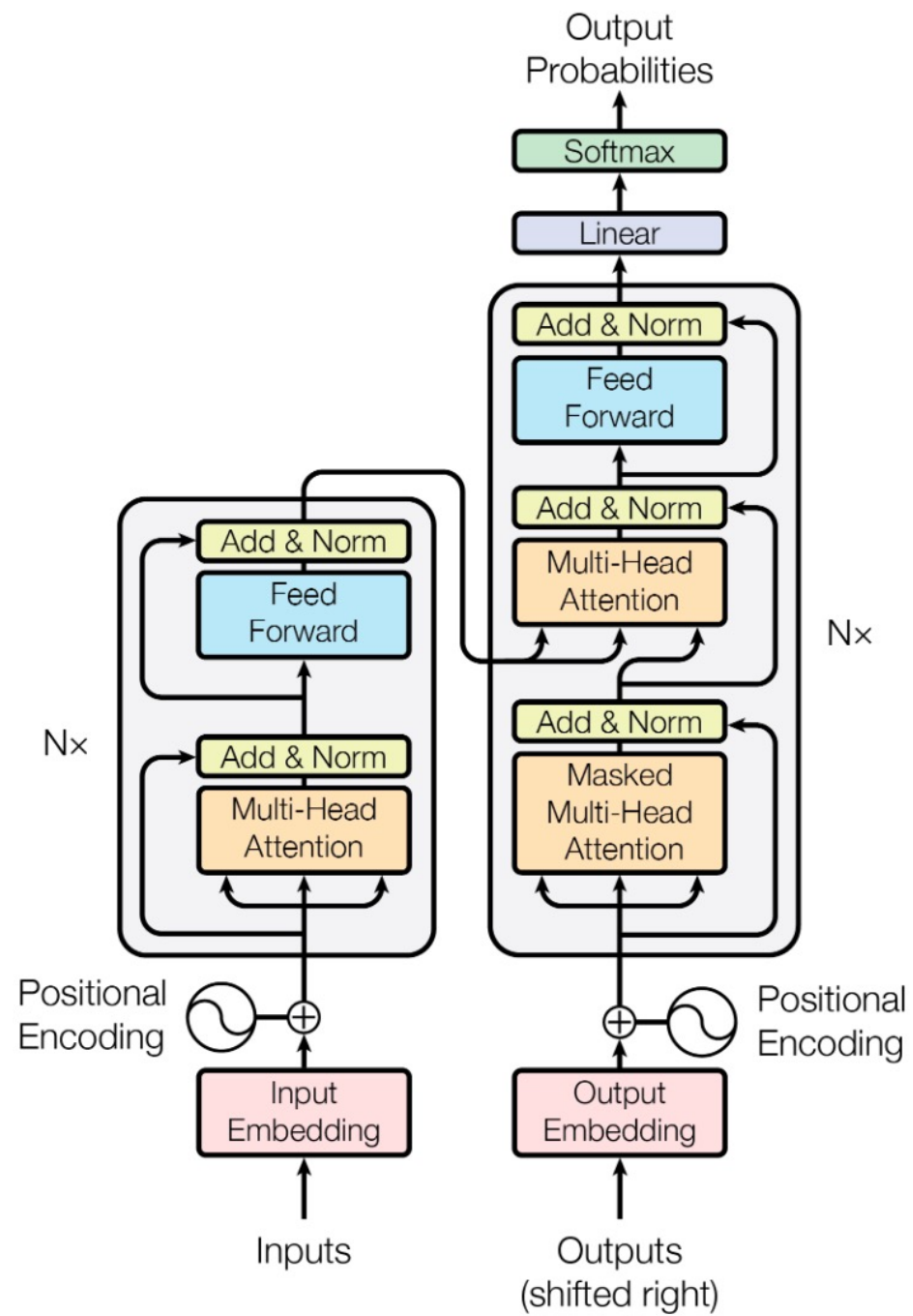


Text Generation with GPT

GPT

- Generative Pre-trained Transformer
- Transformer model that is trained to predict the next word in a sentence
- Trained over many NLP tasks using different model output layers
- GPT -> GPT-2 -> GPT-3 -> GPT-3.5





GPT-3

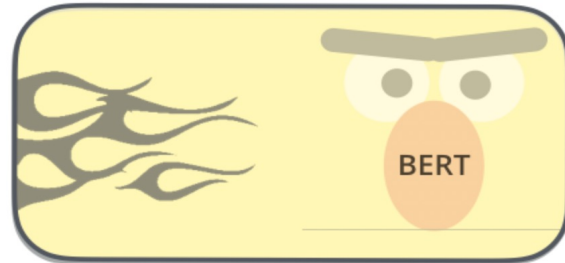
175 Billion Parameter Language Model

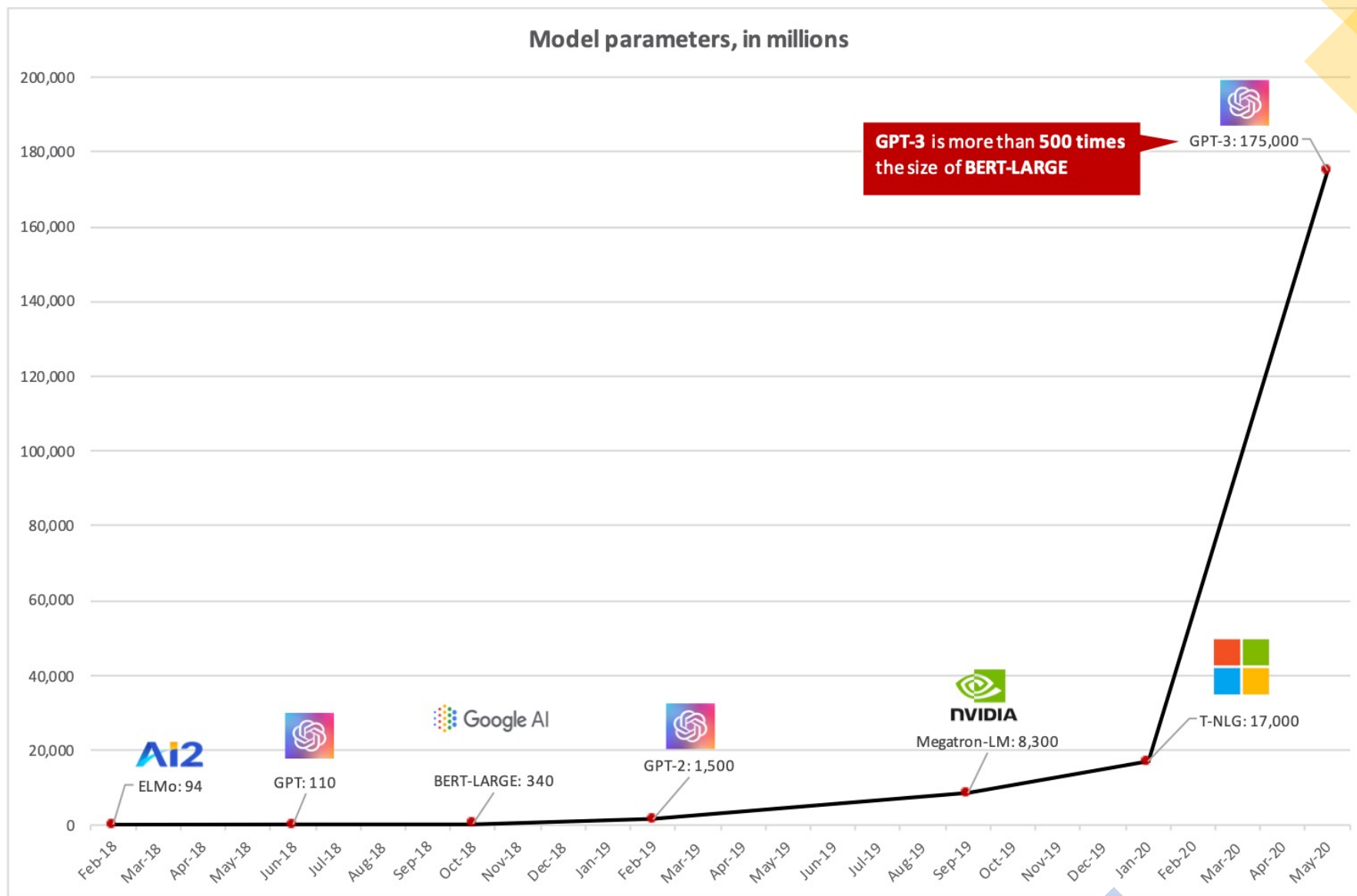
96 layers X 96 attention heads

Trained on a corpus of 300 Billion words

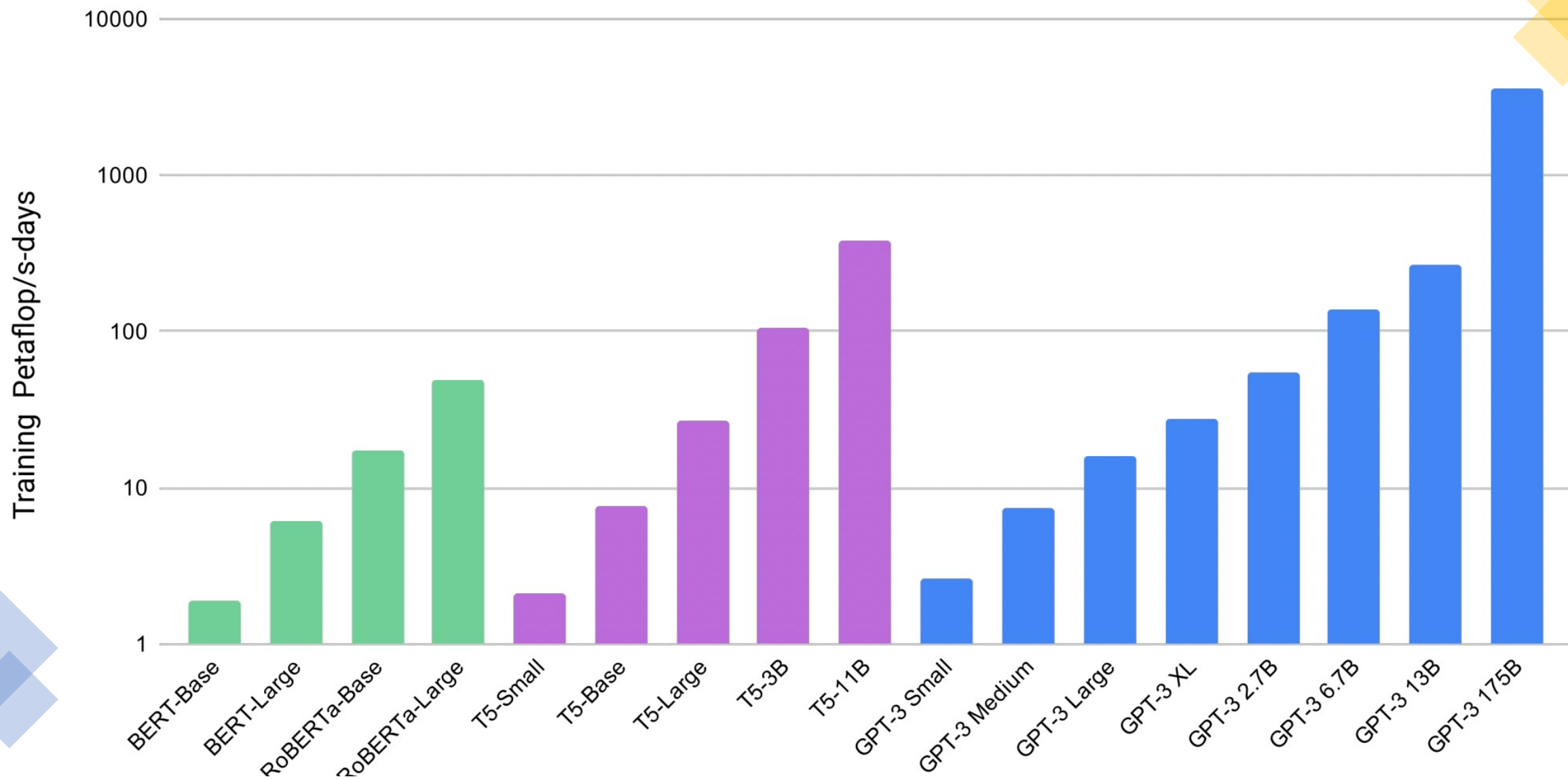
Predicts the next word in a sequence.

Transformer Family

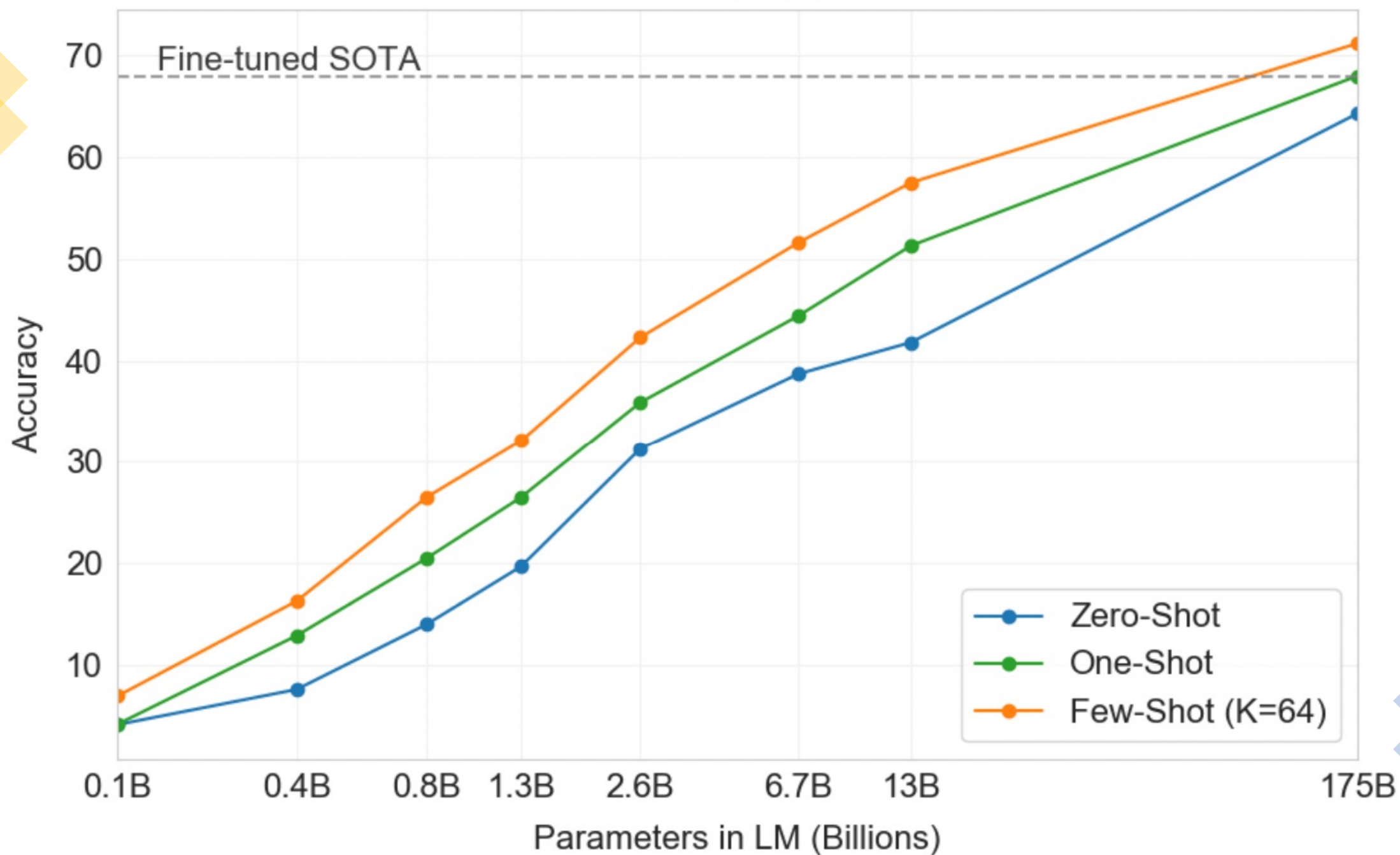




Total Compute Used During Training



TriviaQA



Translation (Multi-BLEU)

