

Fine-tuned GPT2 model

Training process

Unsupervised pre-trained model



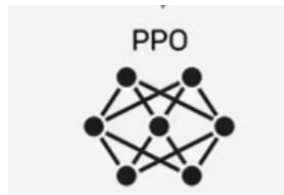
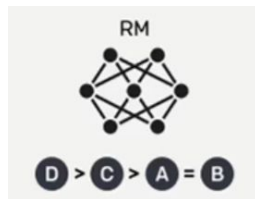
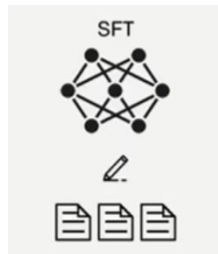
Supervised learning model



Reward model

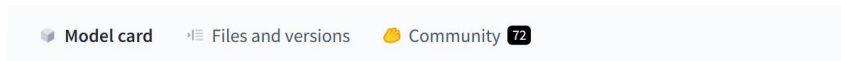


Reinforcement learning model



Unsupervised pre-trained model

- model from hugging face



GPT-2

Test the whole generation capabilities here: <https://transformer.huggingface.co/doc/gpt2-large>

Pretrained model on English language using a causal language modeling (CLM) objective. It was introduced in [this paper](#) and first released at [this page](#).

Disclaimer: The team releasing GPT-2 also wrote a [model card](#) for their model. Content from this model card has been written by the Hugging Face team to complete the information they provided and give specific examples of bias.

- The answer is not useful



Supervised model

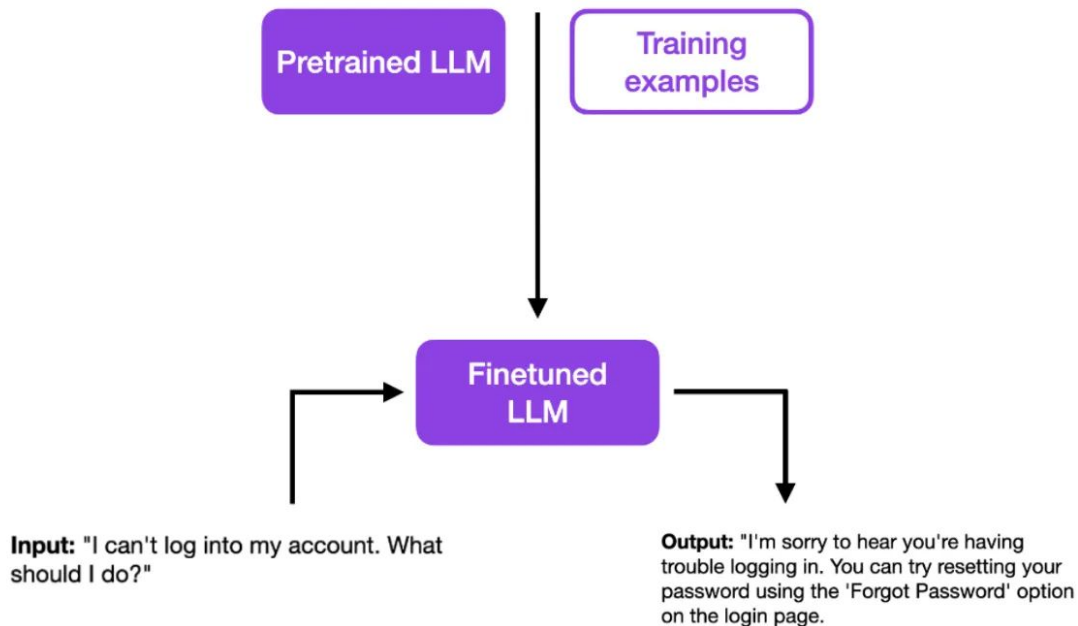
Use our own dataset for fine-tuning training.



Connect the question answer pair and mask answer



Fine-tune the model before training



Future task

- Sort the generated answers to get a new reward model dataset
- Training reward model
- Add reinforcement learning model