Fine-tuned GPT2 model

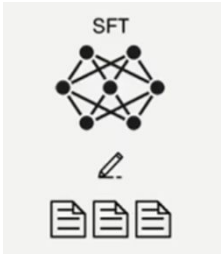
Training process

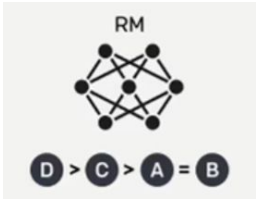
Unsupervised pre-trained model

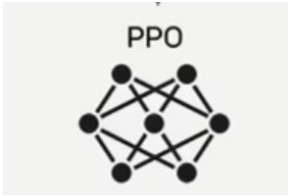
Supervised learning model

Reward model

Reinforcement learning model







Unsupervised pre-trained model

- model from hugging face

Model card | Files and versions | Community 72

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.

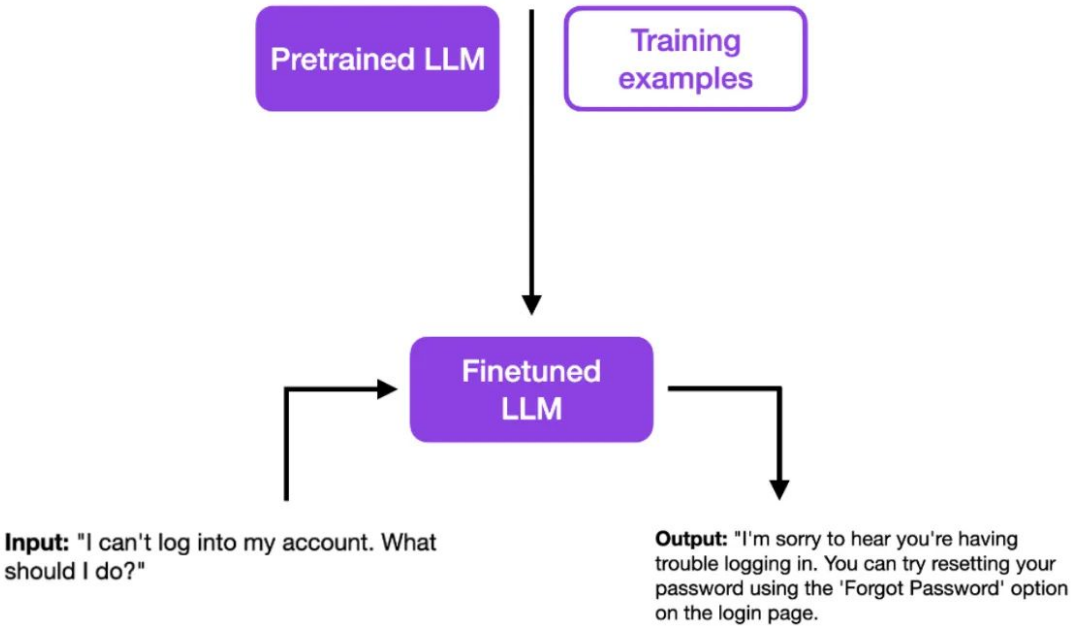
- The answer is not useful

Supervised model

Use our own dataset for fine-tuning training.

Connect the question answer pair and mask answer

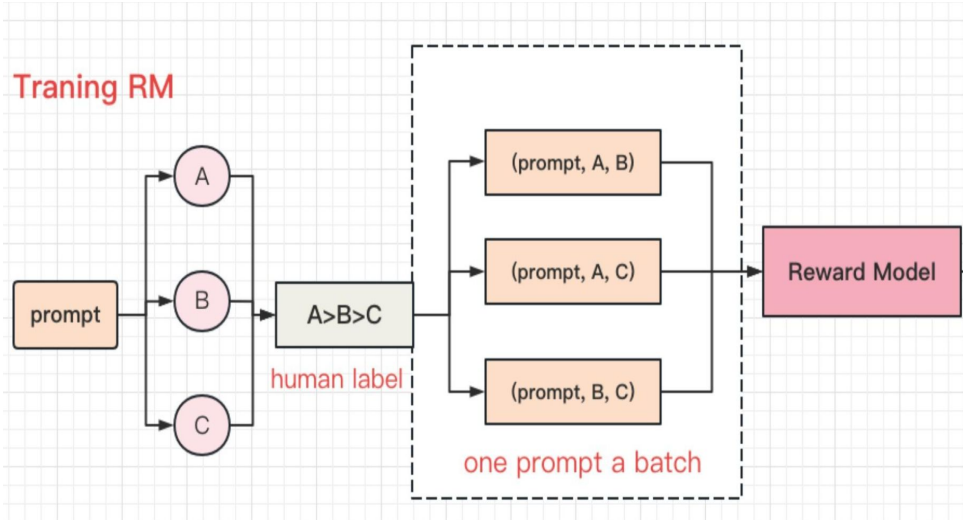
Fine-tune the model before training



Reward Model

- Use the previous fine-tuning model to generate multiple answers, sort the answers to obtain the data set

- The Reward Model is trained on a new data set to rank the answers so that the quality of the answers can be quantified.



Future task

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