

参考資料

ウェブページ

- <https://nlp.seas.harvard.edu/annotated-transformer/>
 - <https://neptune.ai/blog/how-to-code-bert-using-pytorch-tutorial>
 - <https://github.com/huggingface/transformers/tree/main/src/transformers/models/llama>
 - <https://github.com/facebookresearch/llama/tree/57b0eb62de0636e75af471e49e2f1862d908d9d8/llama>
 - <https://colab.research.google.com/drive/1iERDk94Jp0UErsPf7vXyPKeiM4ZJUQ-a?usp=sharing#scrollTo=Cc8354XxCiWI>
 - https://pytorch.org/tutorials/beginner/chatbot_tutorial.html#preparations
 - <https://github.com/lucidrains/PaLM-rlhf-pytorch/tree/main>
 - <https://huggingface.co/blog/rlhf>
 - <https://huggingface.co/learn/deep-rl-course/unit8/hands-on-cleanrl#lets-code-ppo-from-scratch-with-costa-huangs-tutorial>
-
- 論文
 - <https://arxiv.org/pdf/2104.09864.pdf>
 - <https://arxiv.org/pdf/2104.09864v4.pdf>
 - <https://arxiv.org/pdf/2302.13971.pdf>
 - <https://arxiv.org/pdf/1706.03762.pdf>

- 生成方法は llama1 の方法を使った
- データ: 事前学習: pytorch にある WikiText2 fine-tuning
&rl : Cornell Movie-Dialogs
- トレーニング: wikitext2 mask lm で事前学習 Cornell Movie-Dialogs で fine tuning
- Hugging face のライブラリーを使って lora を入れるそのままトレーニング ppo&actorcirtic で強化学習をする
- 結果: chatbot 作るのが失敗した • Llama の rotary position embedding をフーリエ変換に変えた • 生成する時 Causal lm と top p sample で出力した logits を pca で圧縮し転置して一番目のものを選んだ
- 強化学習の reward model では出力の label をコサイン類似度かける 10 でトレーニングした

結果

input

would you like a cup of coffee?

pre_len

12



temp

0.37



クリア

送信

output

called insisted confess which belongny THE
likedjimachi How hang entraps

フラグする

input

if you cooperate with the da maybe they ll help
you with your situation

pre_len6

temp0.62

クリア送信

output

bs fier done ago blind

フラグする

入力するものと出力するものは全然対話になってない

