

# hw3\_training\_lora

姓名: 周睿翔

ID: B11115019

# How I train the model

我先透過助教給的網址，然後從hugging face拉下訓練所需的文件夾，至於訓練步驟，已詳細寫在Readme.md裡面，可以點開看看我訓練的步驟為何。

# 訓練參數

這次訓練, 我個個參數是給

epoch: 200

learning\_rate: 0.0001

learning\_rate調度器: linear(因為一開始訓練速度很慢, 後來查到說, 小型的訓練不需要用cosine, 用linear即可)

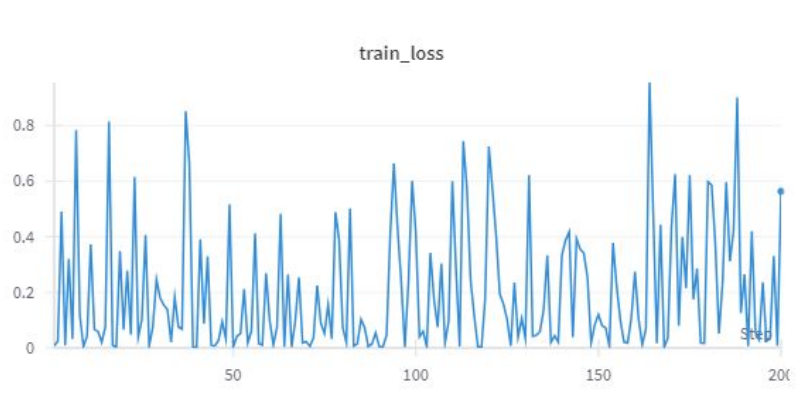
gradient\_accumulation\_steps: 1

checkpointing: 20

其餘參數就是照他的預設值

# wandb訓練曲線

看起來滿糟的我知道，個人認為，是 $r$ 太高，梯度下降時在低點附近跳來跳去，沒有辦法收斂，下次有這種情況，會在調整一些其他參數，使他能夠有辦法收斂



Thank you!!!