

HW10 Report

STuser19 賴昱凱

1. Please explain the training methods in ChatGPT.

Step 1. 文字接龍 next token prediction

讓 model 在網路上抓大量的文本資料，並學習如何在過去的所有輸入文字後接上合理的文字內容，例如之前訓練 transformer decoder 的方式。而且本步驟**不需要人工標注內容**，因此可以輕鬆的用十分大量的資料做訓練，讓 model 有一定的語言能力。

Step 2. 人類引導文字接龍的方向

這個步驟就是由人工標注的內容做訓練，藉由人類生成的問題及在其後接上解答，藉此訓練模型知道人類偏好的回答。此步驟不需要極大量資料來訓練，只需要讓模型知道人類喜好即可。

Step 3. 模仿人類的喜好

藉由輸出多種的答案，由人類評斷出各答案的優劣，並生成一個模型使自己的輸出以人類的評斷可以獲得高分，學習到人類偏好的回答方式。

Step 4. 用增強式學習模擬人類學習

現在已經有了可以評分結果的模型，我們將 GPT 模型的輸出拿去輸入評分模型，並藉由調整參數提高 Reward，最終訓練完成就可以獲得成熟的 ChatGPT。

2. Please explain why do we mostly use fine-tuning instead of training from scratch on language models?

i. 運算資源與時間成本

從頭訓練一個大型模型所耗費的運算量及時間都不是一般實驗室可以承受的，因此在已經有一定能力的模型上作為調才是更有效率、划算的訓練方式。

ii. 專注任務內容

Fine-tuning 可以讓模型專注在需要學會的任務內容上，不需要學會例如對話、文法等基本知識，更可以有效率的訓練模型。

iii. 小量數據的高效率運用

往往大多專門的任務資料量都不大，若使用小量資料對模型從頭訓練，會容易出現 overfitting 的情況，在 fine-tuning 上就比較不容易出現此問題。

Explain what is instruction fine-tuning?

Instruction fine-tuning 是一種 fine-tuning，可以讓模型更在某個領域變得更加專業，也可以更好的回答相關問題。大多的 pre-train LLM 都只有 next token prediction 的能力，對於接收指示並依照指示回答的能力都不太好，因此利用 instruction fine-tuning 訓練其遇到指示做回答的能力。

訓練方式其實不複雜，就是輸入指令並預期輸出結果，與先前提過的 supervised learning 基本相同，只是因為他是在 pre-train model 上訓練，因此不需要非常大量的資料量，就可以讓模型學會如何根據指示執行任務。