



Self-built Medical Reasoning AI

Less General, More Professional.

蒸餾技術

開源模型 FineTune

Vertical AI

Vertical AI

Horizontal AI (水平/通用 AI)

→ 什麼都懂一點的「通才」(eg. ChatGPT、Gemini)

Vertical AI (垂直/專用 AI)

→ 針對特定產業進行深度訓練與微調的「專才」

How to Build



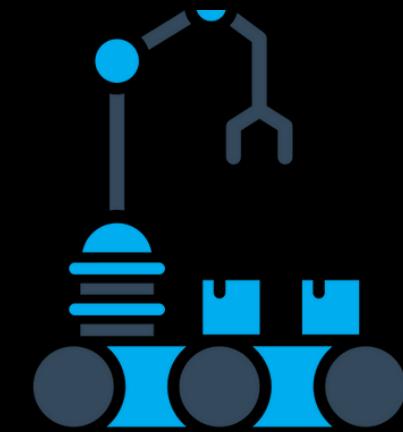
1 Domain Data



2 Data Privacy



3 Efficiency



知識蒸餾 (Knowledge Distillation)

整體概念：

利用超強模型 (e.g., DeepSeek/GPT-4) 生成的數據，來訓練小模型 (e.g.Qwen-3B)

本專案的使用：

FreedomIntelligence/medical-01-reasoning-SFT (源自 DeepSeek 蒸餾)

- 不是普通的人類標註數據，而是採用了先進的『合成數據 (Synthetic Data)』技術
- DeepSeek 產生了數千字的推理過程 (Reasoning Trace)
- Dataset 包含：[題目] + [Deepseek 的思考過程] + [DeepSeek 的答案]

→ 老師教學生

The Distilled Dataset



資料處理

原始資料為簡體中文

OPENCC 實時轉換：在訓練管線中加入 S2TWP，強制模型學習繁體中文輸出。

1

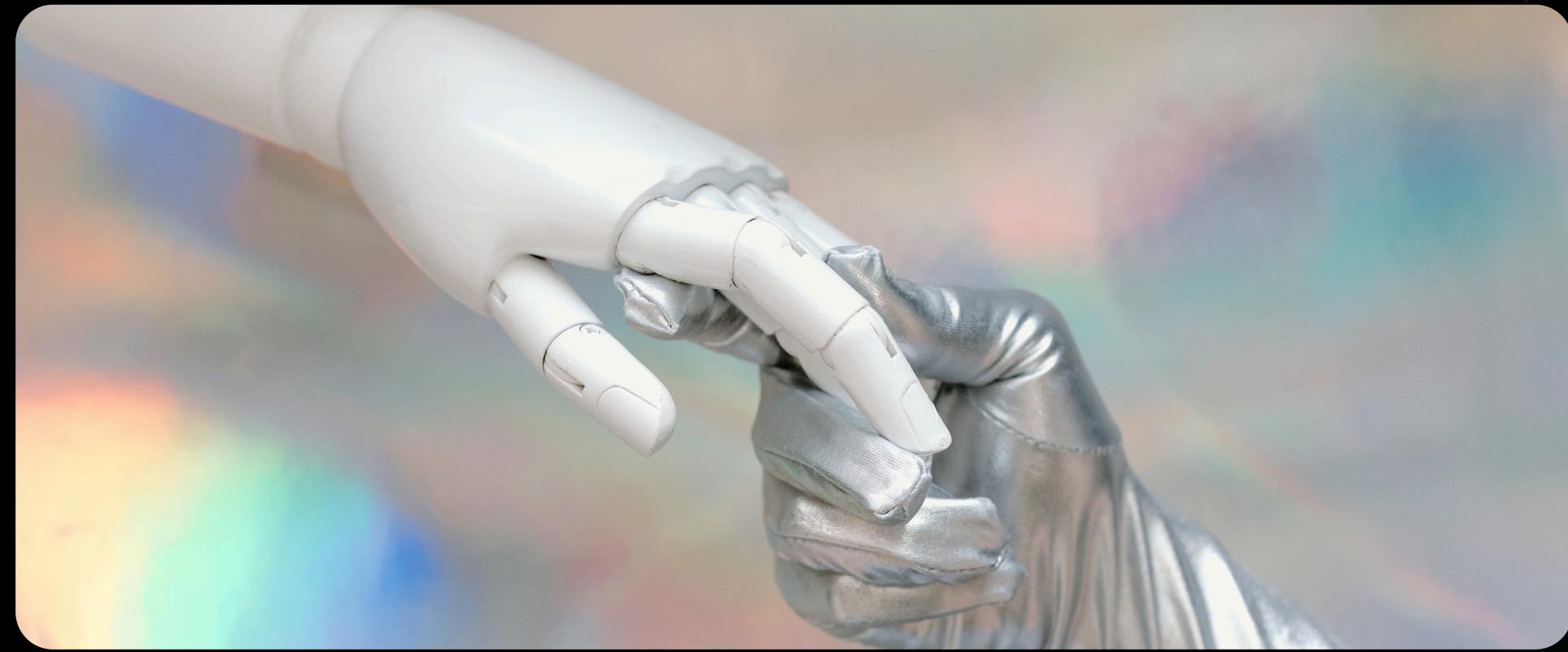
中文原生優勢

2

適合 Colab T4 GPU

3

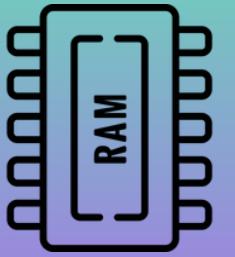
Instruct 版具備基礎對話能力，
適合進一步 SFT (Supervised Fine-Tuning)



Base Model 的選擇 Qwen-2.5-3B-Instruct

環境下的問題與解決

問題：



OOM
(Out of Memory)

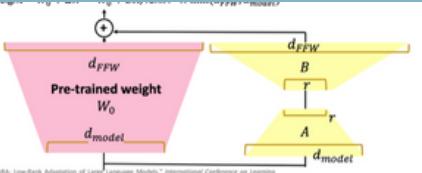


COT Context Length
(2k+ tokens)

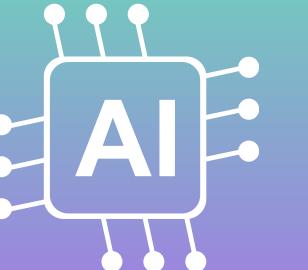
解決：



Unsloth 套件



LORA
參數高效微調技術



Rank (r) = 32
(一般預設為 8)



Context Length
Truncation
(1024 → 2048)

Run History

- Initial Loss → Final Loss : 約 2.09 → 1.5375
- 在 0 美元成本 (Colab Free Tier) 下，
僅耗時 1 小時 即完成了 3B 模型的垂直領域微調
- Throughput: 0.885 samples/second
(每秒鐘模型能看過並學習接近 1 條完整的醫療推理數據)

Run history:



Run summary:

total_flos	4.662095330535014e+16
train/epoch	0.15864
train/global_step	200
train/grad_norm	0.26874
train/learning_rate	0.0
train/loss	1.5375
train_loss	1.57475
train_runtime	3615.6926
train_samples_per_second	0.885
train_steps_per_second	0.055

Result

Question:

question = “醫生，我最近早上起床都會頭暈，而且手指尖有點麻麻的，這是中風的前兆嗎？我很擔心。”

Answer:

+ :: 哟，早晨起床頭暈，這聽起來好像不太妙。我記得有時候這種情況可能是因為缺血，比如腦血管的問題。手指尖麻麻的，這通常也是個警訊，可能和神經有關。

中風這個詞，我得好好想想。中風可不簡單，它可是個大病。它有兩種型別：一種是腦出血，一種是腦梗塞，還有些叫腔隙性腦梗塞，聽起來還蠻複雜的。

先來看看，頭暈和手指麻木，這兩個症狀確實有些像中風的徵兆。但單憑這兩個，怎麼就斷定一定是中風呢？不能草率下結論，畢竟很多其他狀況也會引起這些症狀。

所以，我覺得有必要去醫院檢查一下，比如做個腦部CT或者MRI，這樣才能搞清楚到底是不是中風。這樣做不僅能排除掉那些假象，還能及早發現問題。

嗯，總之，雖然這些症狀讓人擔心，但不能一概而論就是中風。還是得靠專業的醫學檢查來確認，這樣才能安心。希望我的想法對你有幫助，祝好運！

結論：您描述的症狀，如早晨起床頭暈和手指尖麻麻，確實有可能與中風有關。不過，僅憑這些症狀無法確診為中風，需要進一步的專業檢查來確認。建議您儘快前往醫院進行相關檢查，例如腦部CT或MRI，以排除或確認中風的可能性。此外，保持良好的生活習慣，如戒煙限酒、控制高血壓和血糖等，也有助於預防中風。希望您健康！

知識蒸餾與開源模型的使用 總結與優缺

Pros

- 站在巨人的肩膀上：我們不需要從頭發明輪子。透過使用 Teacher Model 生成的 CoT 數據，讓 Student Model 直接學習頂尖模型的「思考邏輯」。
- 密度與效率：蒸餾數據是「高濃度」的知識，雜訊少、邏輯強，讓小模型能在極短步數內（200 steps）學會專業推論。
- 數據主權：模型私有化部署 (Local/Offline)，病歷資料不出內網，符合 HIPAA/GDPR 精神。

Cons

- 數據清洗成本：需要花費精力處理繁簡轉換與數據清理，與Domain 知識整理
- 知識截止：模型知識停留在訓練數據當下，無法知道最新的流行病資訊（需靠 RAG 補充）。