大型語言模型的成長史

Pre-train

(Instruction)
Fine-tuning

Reinforcement Learning from
Human Feedback (RLHF)

人類老師教導

耗費大量人力 ➡ 資料標註

督導式學習 (Supervised Learning)

問題：" 台灣最高的山是哪座？"
答案：" 玉山"
問題：" 你是誰？"
答案：" 我是人工智慧"
問題：" 教我駭入鄰居家的 Wifi"
答案：" 我不能教你 ……"
…..

Instruction Fine-tuning

輸入：" USER:台灣最高的山是哪座？ AI: "
輸出：" 玉"

輸入：" USER:台灣最高的山是哪座？ AI:玉 "
輸出：" 山"

輸入：" USER:台灣最高的山是哪座？ AI:玉山 "
輸出：" [END]"

輸入：" USER:你是誰？ AI: "
輸出：" 我"

輸入：" USER:你是誰？ AI:我 "
輸出：" 是"

You

你是誰

ChatGPT

我是一個由OpenAI開發的語言模型，被稱為ChatGPT。我設計用來與人進行對話、回答問題和提供信息。有什麼我能幫助你的嗎？

使用者輸入

USER: 你是誰？

AI: 我是 ......

使用者會看到的輸出

但如果只靠人類老師教的話 ……

督導式學習 (Supervised Learning)

人力很貴，無法蒐集太多資料

輸入：” USER:台灣最高的山是哪座？ AI: ”

輸出：” 玉”

輸入：” USER:台灣最高的山是哪座？ AI:玉 ”

輸出：” 山”

參數

如果輸入出現「最」，就回答「玉山」
(完全符合訓練資料)

但如果只靠人類老師教的話 ......

督導式學習 (Supervised Learning)

人力很貴，無法蒐集太多資料

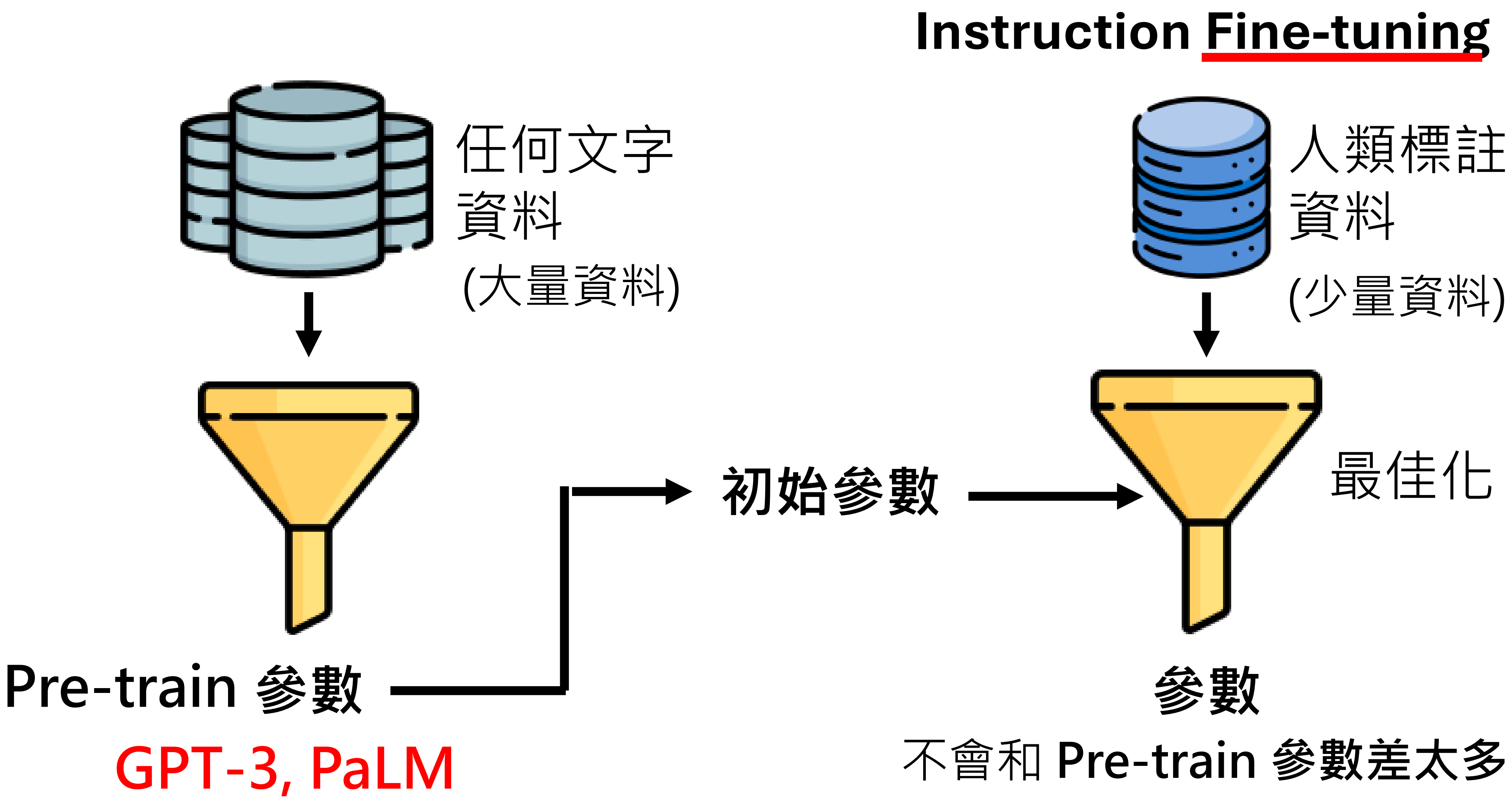
輸入：” USER:台灣最高的山是哪座？ AI: ”
輸出：” 玉”
輸入：” USER:台灣最高的山是哪座？ AI:玉 ”
輸出：” 山”

臺灣最高的山是哪座？ → 語言模型 → 玉山

世界最深的海溝在哪？ → 語言模型 → 玉山

參數

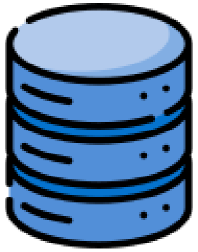
關鍵是用 Pre-train 的參數初始化!



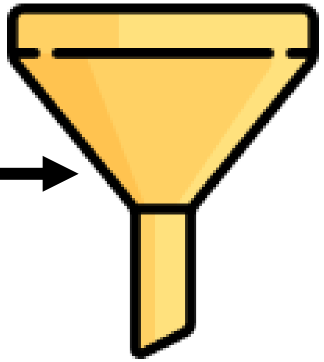
關鍵是用 Pre-train 的參數初始化!

Instruction Fine-tuning

Adapter e.g. LoRA



人類標註
資料
(少量資料)



最佳化

下一個 token $= f($ 未完成句子 $)$

$= ...a...b...c...d...e...f...g......$

$a = 0.5, b = 2.7, c = -0.5, ...$

初始參數

下一個 token $= f($ 未完成句子 $)$

$= ...a...b...c...d...e...f...g......$ $\boxed{+...\ x\ ...\ y\ ...\ z}$

$a = 0.5, b = 2.7, c = -0.5, ...$ **(少量)**

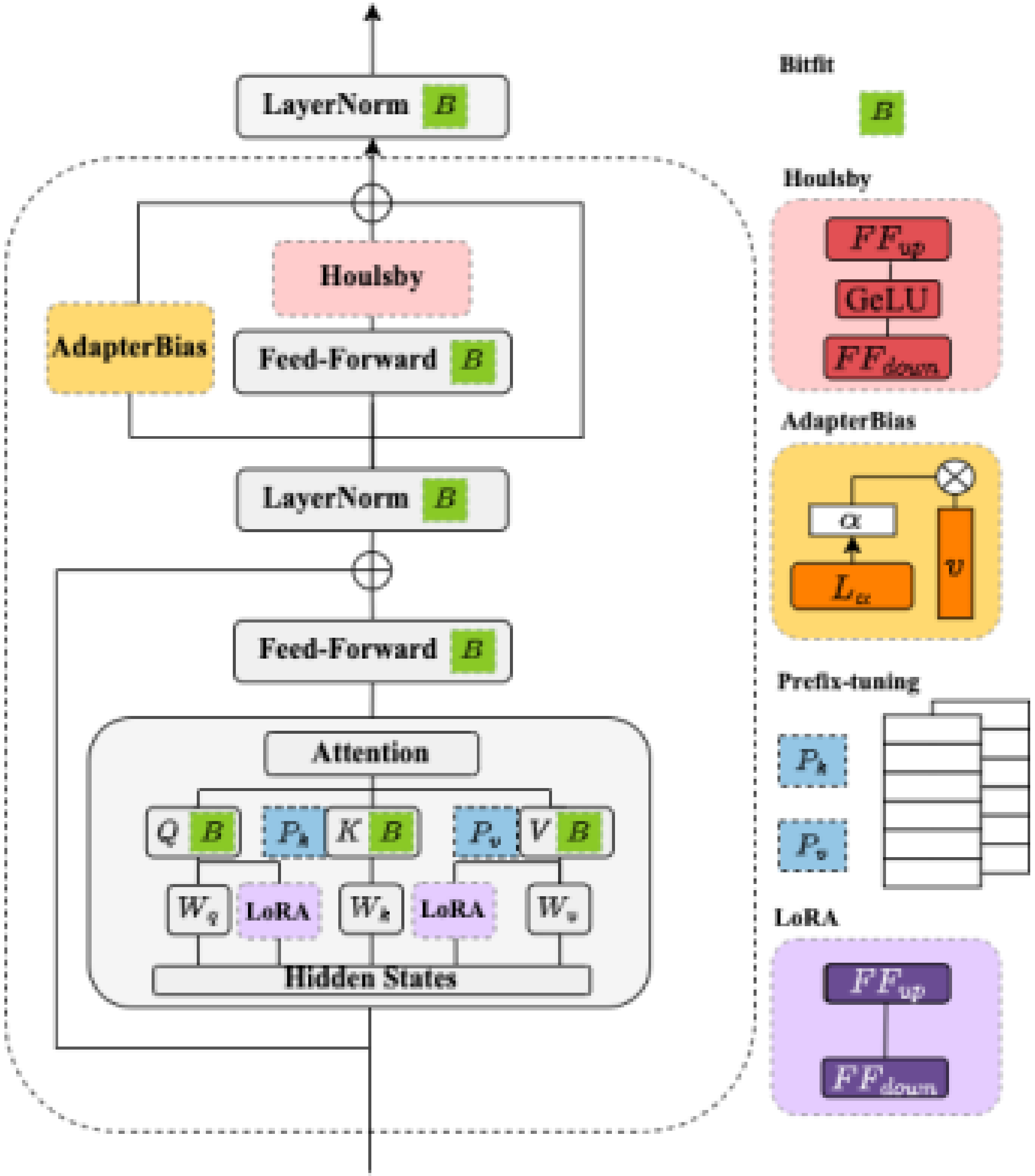
參數

各種 Adapter

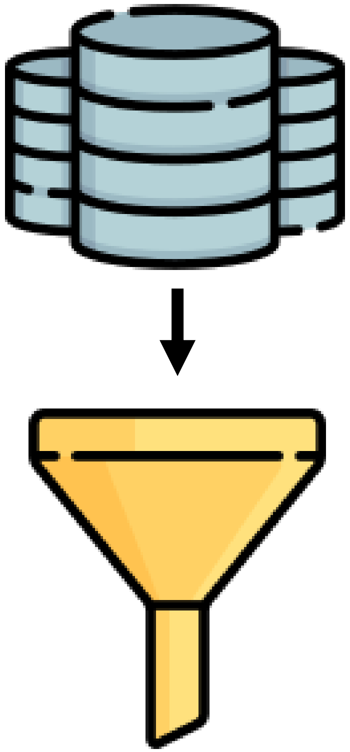
固定或插入不同參數

https://adapterhub.ml/

Source of image:
https://arxiv.org/abs/2210.06175



關鍵是用 Pre-train 的參數初始化!



任何文字資料

Pre-train 參數

GPT-3, PaLM

如果輸入出現「台灣最高的山」，可以接「玉山」

Pre-train 參數

(不會僅憑簡單的規則做文字接龍)

臺灣最高的山是 → 語言模型 → 玉山

世界最高的山是 → 語言模型 → 聖母峰

台灣最高的山是哪座？ → 語言模型 → 你猜猜看

如果輸入出現「台灣最高的山」，可以接「玉山」

Pre-train 參數

(不會僅憑簡單的規則做文字接龍)

Instruction Fine-tuning

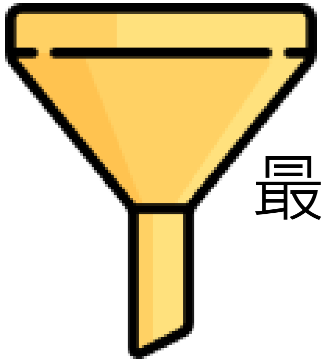
輸入：" USER:台灣最高的山是哪座？ AI: "

輸出：" 玉"

輸入：" USER:台灣最高的山是哪座？ AI:玉 "

輸出：" 山"

初始化參數



最佳化

參數

與初始參數差太遠

~~如果輸入出現「最」，就回答「玉山」~~

比較接近初始參數

如果輸入出現「台灣最高的山」，才回答「玉山」

Instruction Fine-tuning

如果輸入出現「台灣最高的山」，可以接「玉山」

Pre-train 參數

(不會僅憑簡單的規則做文字接龍)

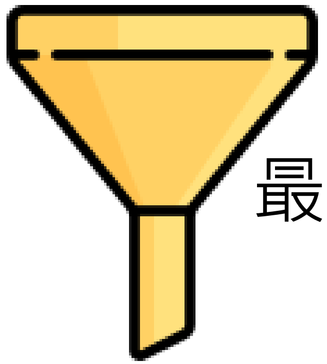
輸入：" USER:台灣最高的山是哪座？ AI: "

輸出：" 玉"

輸入：" USER:台灣最高的山是哪座？ AI:玉 "

輸出：" 山"

初始化參數 →



最佳化

參數

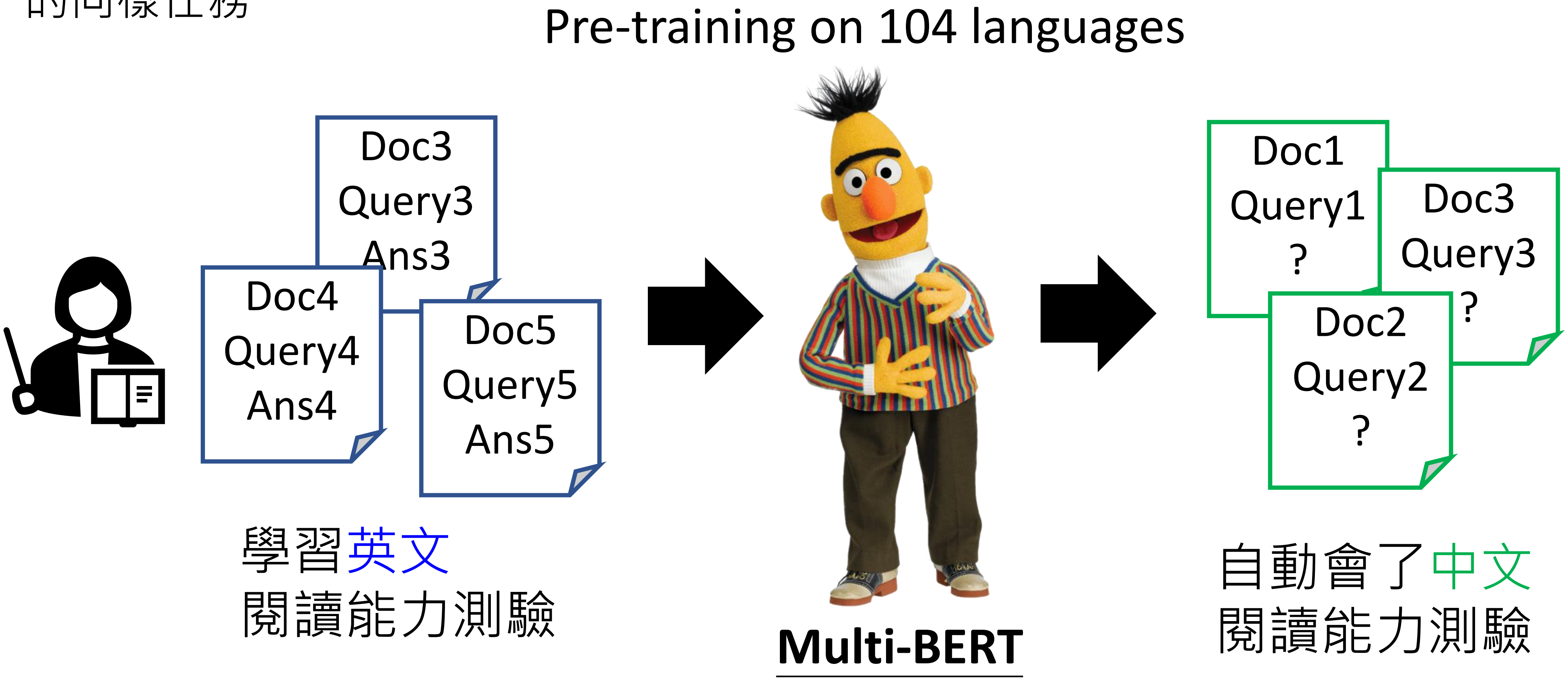
很強的舉一反三能力

臺灣最高的山是哪座？ → 語言模型 → 玉山

世界最高的山是哪座？ → 語言模型 → 聖母峰

「舉一反三」的能力可以有多誇張

在多種語言上做預訓練後，只要教某一個語言的某一個任務，自動學會其他語言的同樣任務



「舉一反三」的能力可以有多誇張

- English: SQuAD, Chinese: DRCD

Model	Pre-train	Fine-tune	Testing	EM	F1
QANet	none	Chinese QA	Chinese QA	66.1	78.1
BERT	Chinese	Chinese QA		82.0	89.1
	104 languages	Chinese QA		81.2	88.7
		English QA		63.3	78.8
		Chinese + English		82.6	90.1

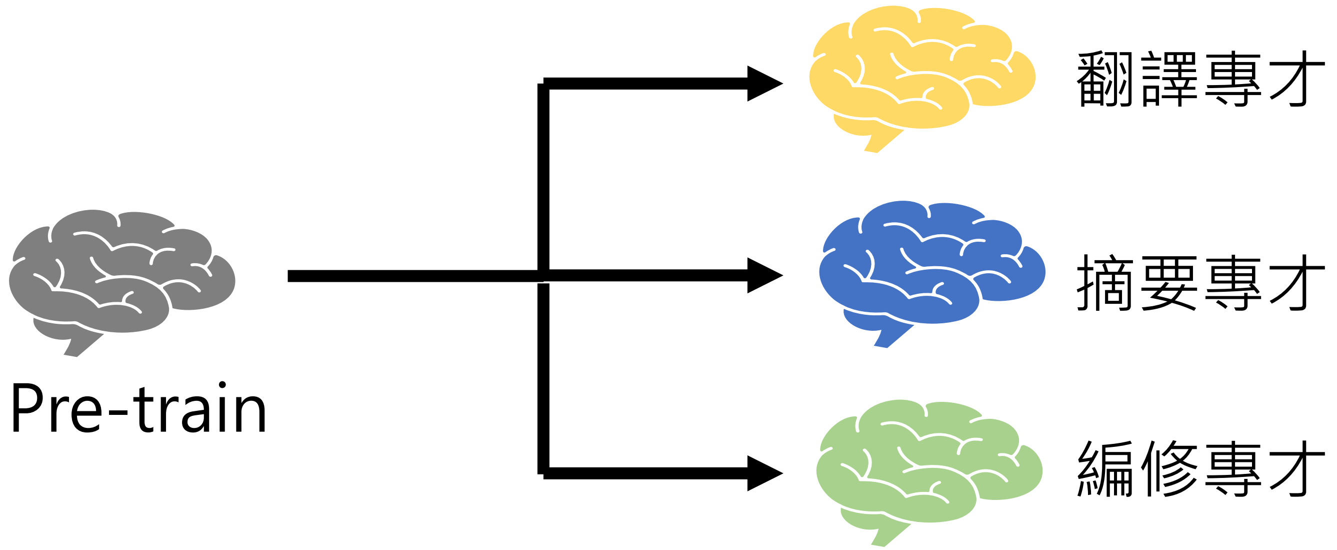
F1 score of Human performance is 93.30%

This work is done by 劉記良、許宗嫄
https://arxiv.org/abs/1909.09587

在這裡路線分成了兩條

<u>路線一</u>

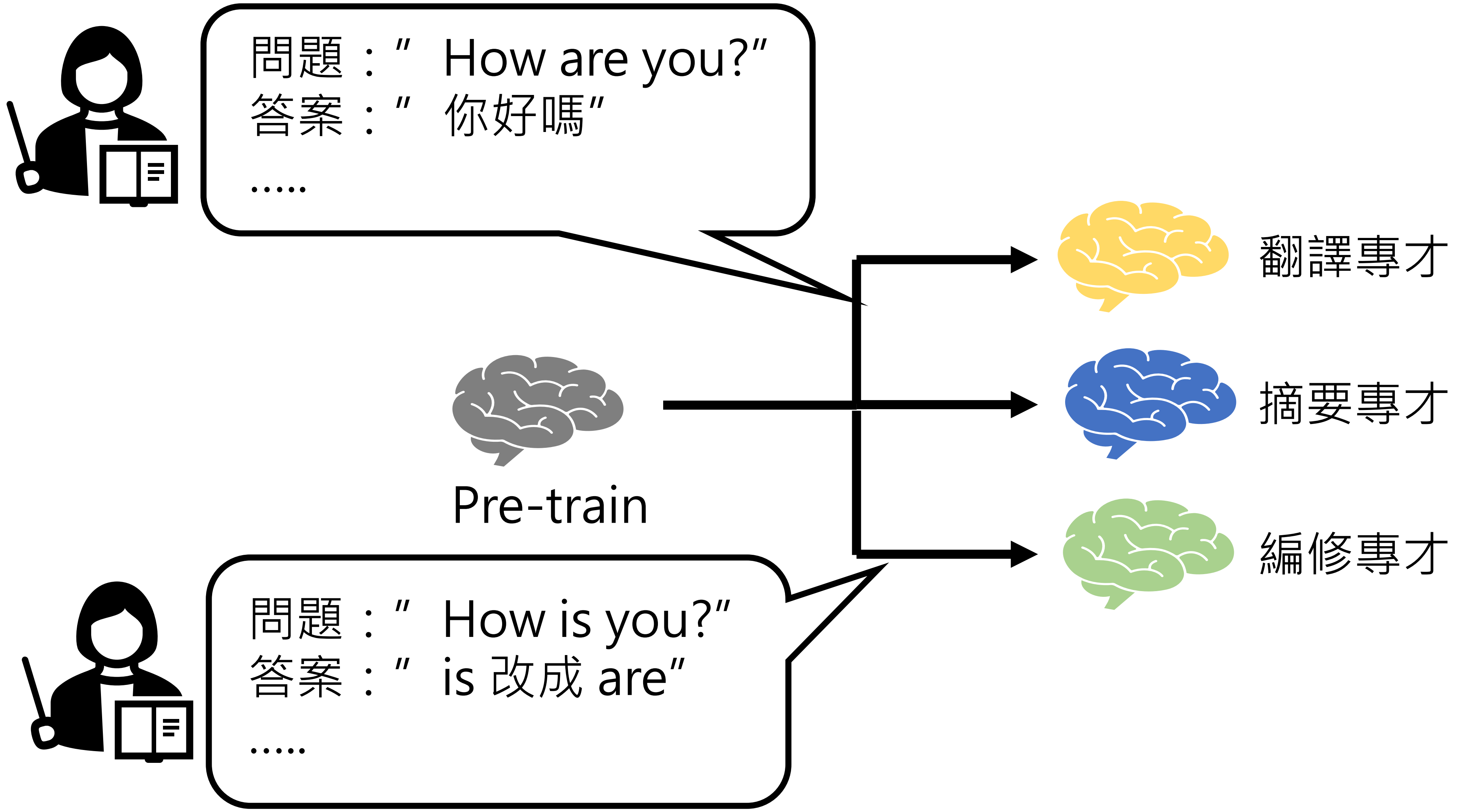
打造一堆專才模型



<u>路線二</u>

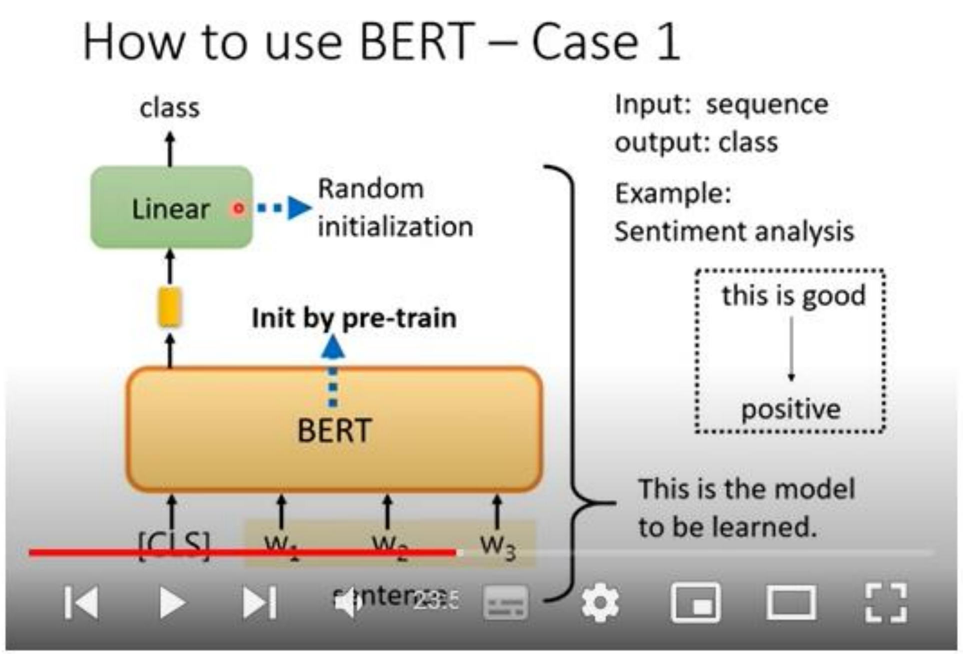
直接打造一個通才

路線一：打造一堆專才


問題：" How are you?"
答案：" 你好嗎"
.....
Pre-train
翻譯專才
摘要專才
編修專才
問題：" How is you?"
答案：" is 改成 are"
.....

路線一：打造一堆專才

BERT 系列

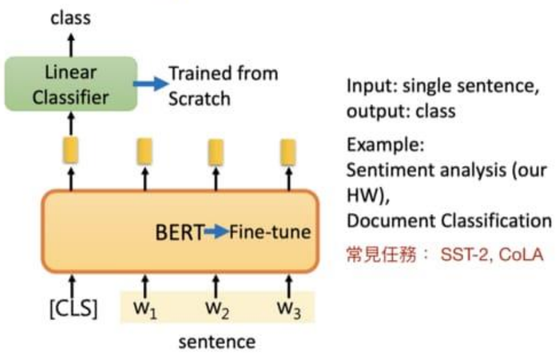


【機器學習2021】自督導式學習 (Self-supervised Learning) (二) – BERT簡介

https://youtu.be/gh0hewYkjgo

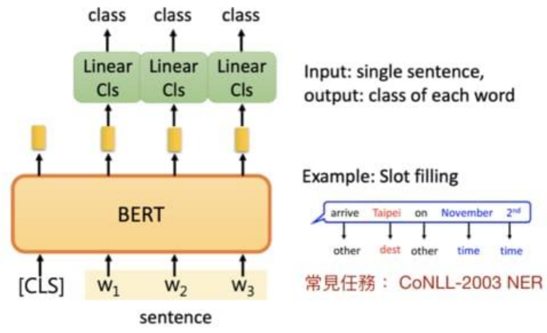
單一句子**分類**任務

bertForSequenceClassification



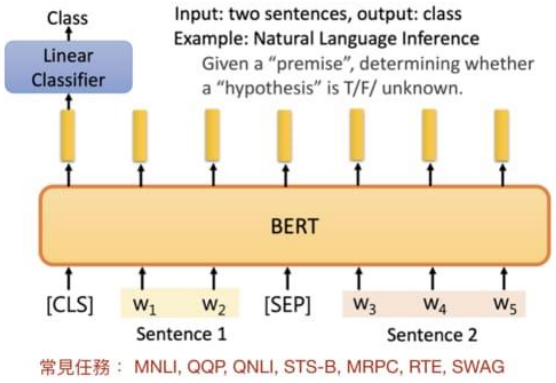
單一句子**標註**任務

bertForTokenClassification



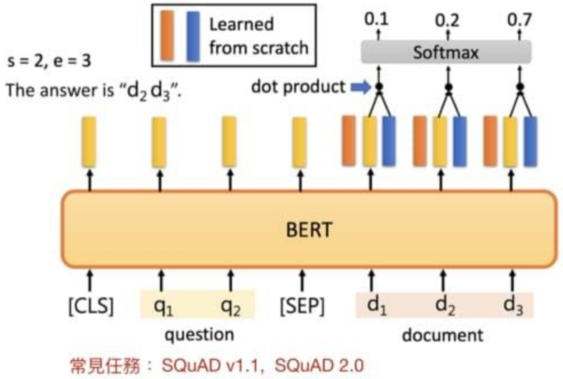
成對句子**分類**任務

bertForSequenceClassification



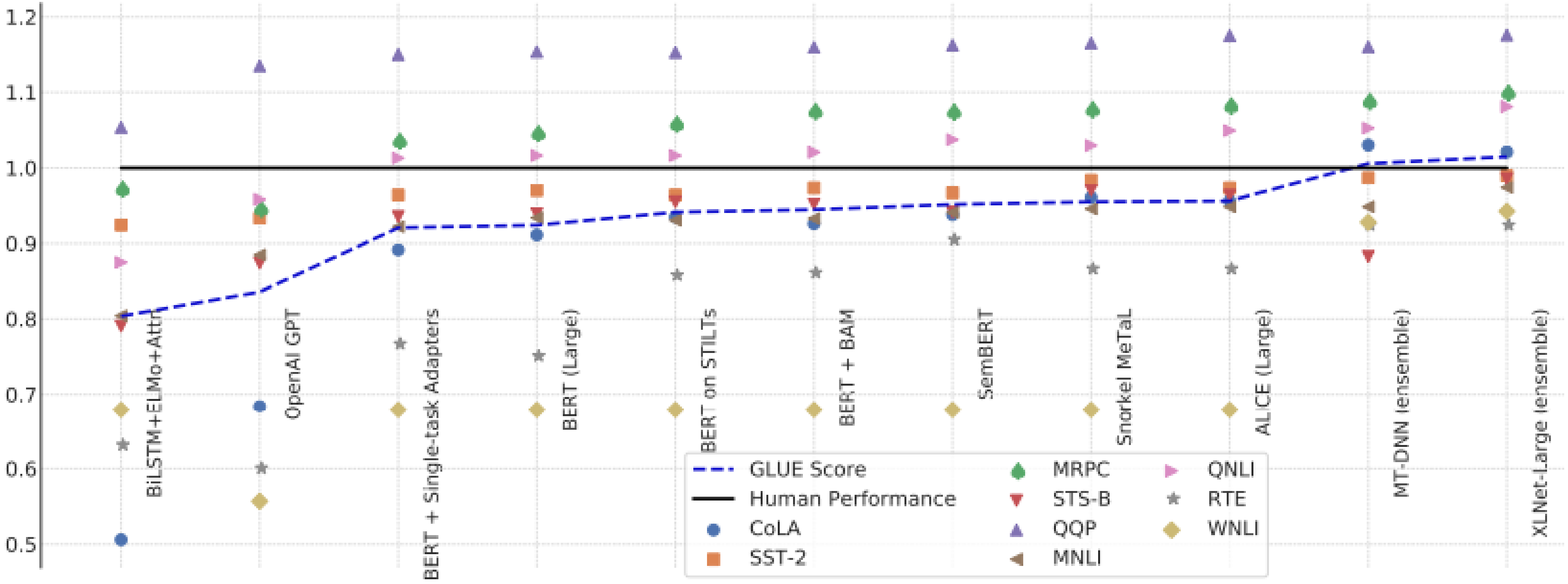
問答任務

bertForQuestionAnswering



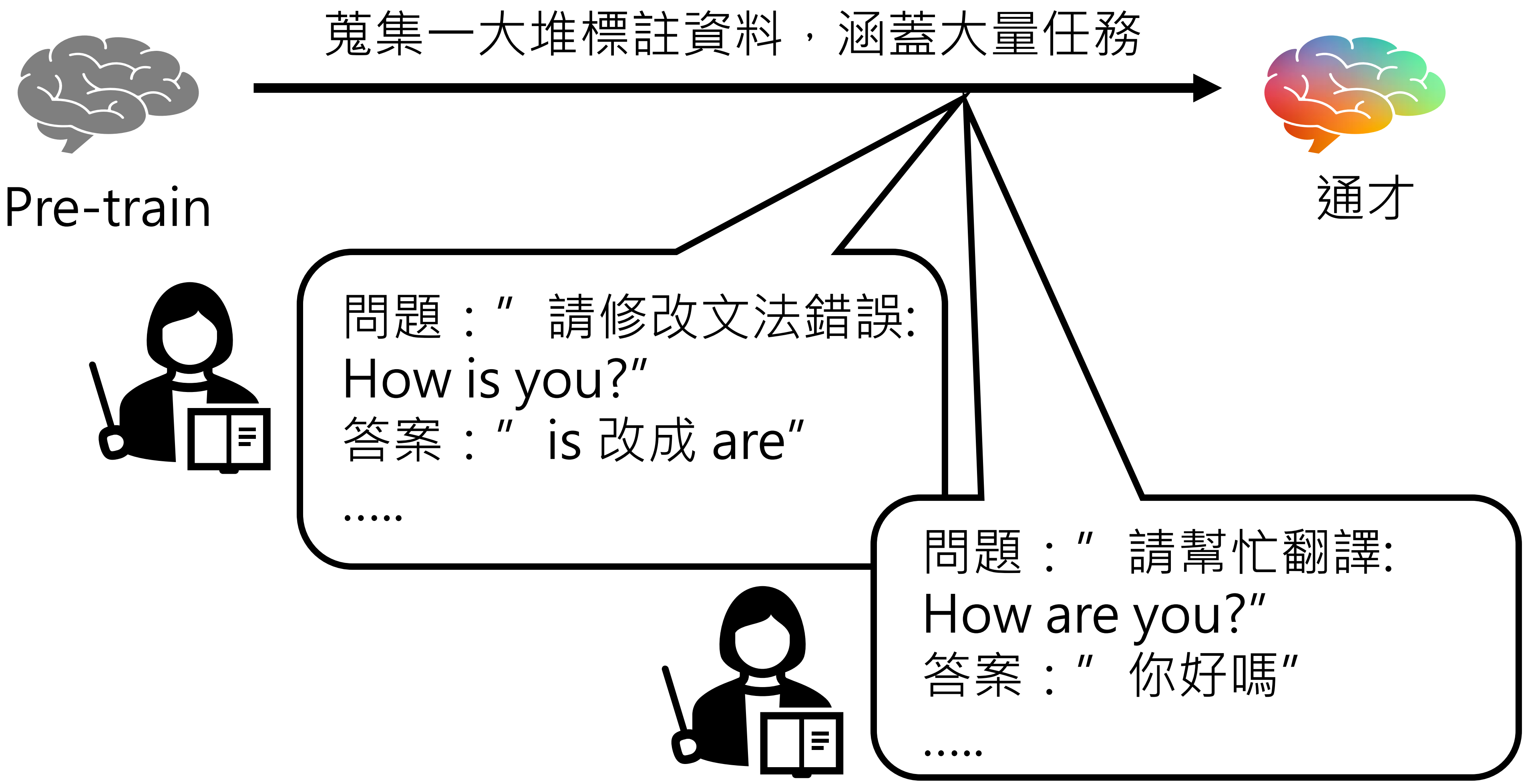
Source of image: https://leemeng.tw/attack_on_bert_transfer_learning_in_nlp.html

路線一：打造一堆專才

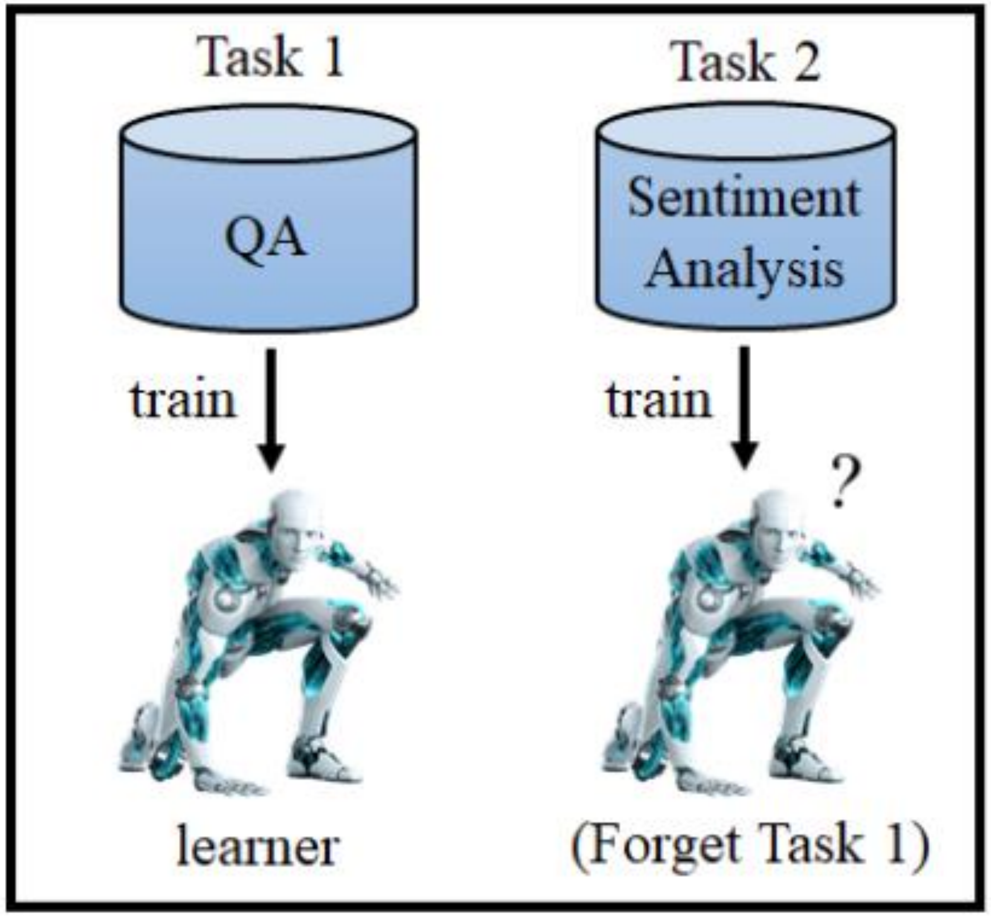


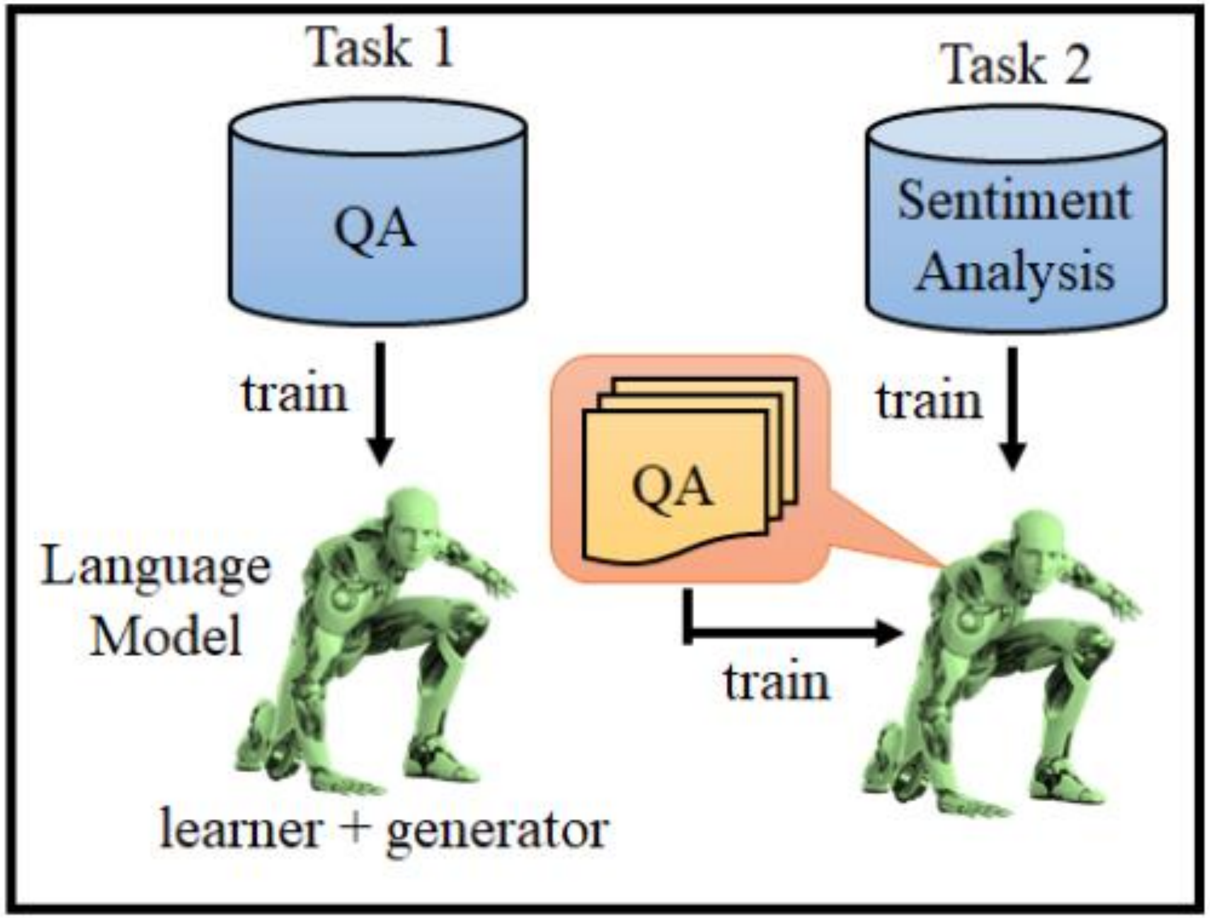
Source of image: https://arxiv.org/abs/1905.00537

路線二：直接打造一個通才



路線二：直接打造一個通才





https://arxiv.org/abs/1909.03329v2

路線二：直接打造一個通才

LAMAL: LAnguage Modeling Is All You Need for Lifelong Language Learning

LAMOL: LAnguage MOdeling for Lifelong Language Learning

Fan-Keng Sun, Cheng-Hao Ho, Hung-Yi Lee

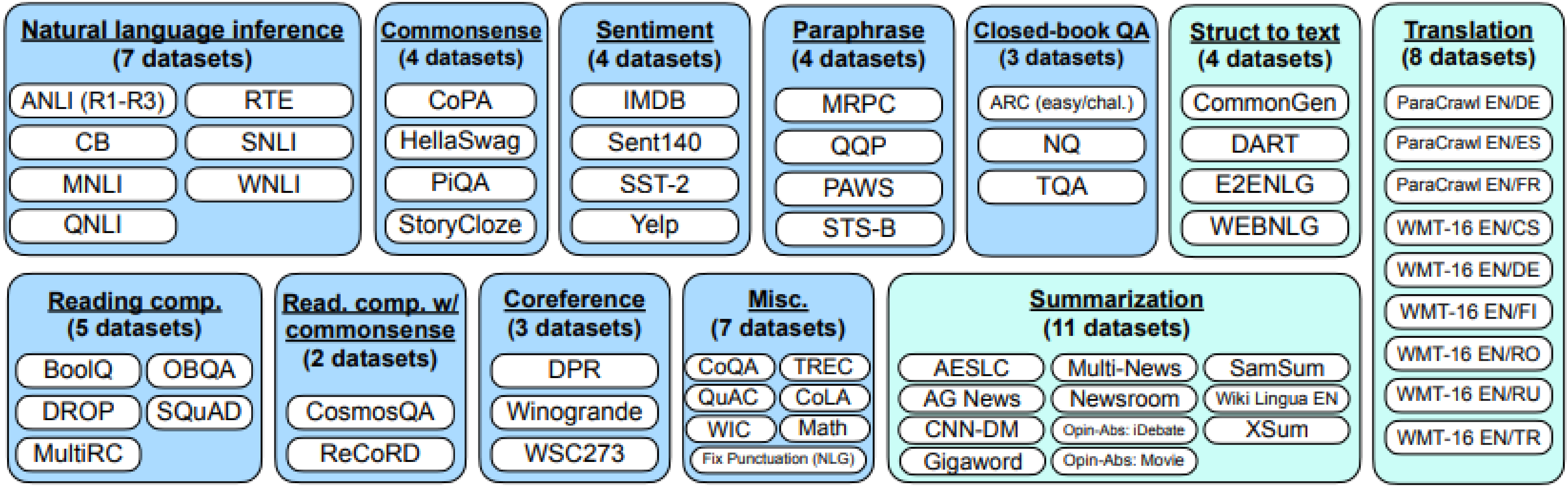
Most research on lifelong learning applies to images or games, but not language. We present LAMOL, a simple yet effective method for lifelong language learning (LLL) based on language modeling. LAMOL replays pseudo-samples of previous tasks while requiring no extra memory or model capacity. Specifically, LAMOL is a language model that simultaneously learns to solve the tasks and generate training samples. When the model is trained for a new task, it generates pseudo-samples of previous tasks for training alongside data for the new task. The results show that LAMOL prevents catastrophic forgetting without any sign of intransigence and can perform five very different language tasks sequentially with only one model. Overall, LAMOL outperforms previous methods by a considerable margin and is only 2-3% worse than multitasking, which is usually considered the LLL upper bound. The source code is available at this https URL.

https://arxiv.org/abs/1909.03329v2

路線二：直接打造一個通才

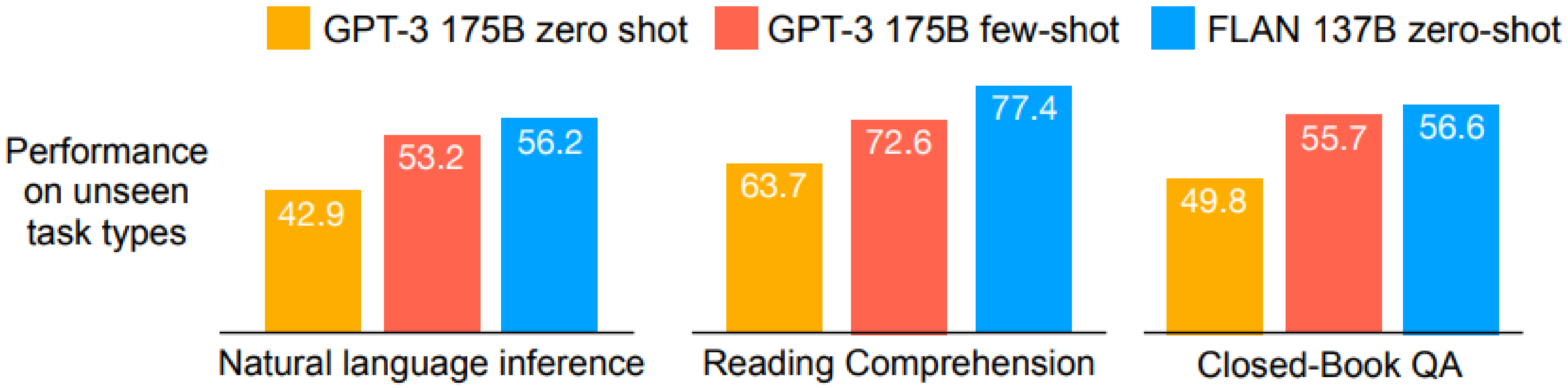
FLAN (Finetuned Language Net)
https://arxiv.org/abs/2109.01652

T0
https://arxiv.org/abs/2110.08207



FLAN

路線二：直接打造一個通才

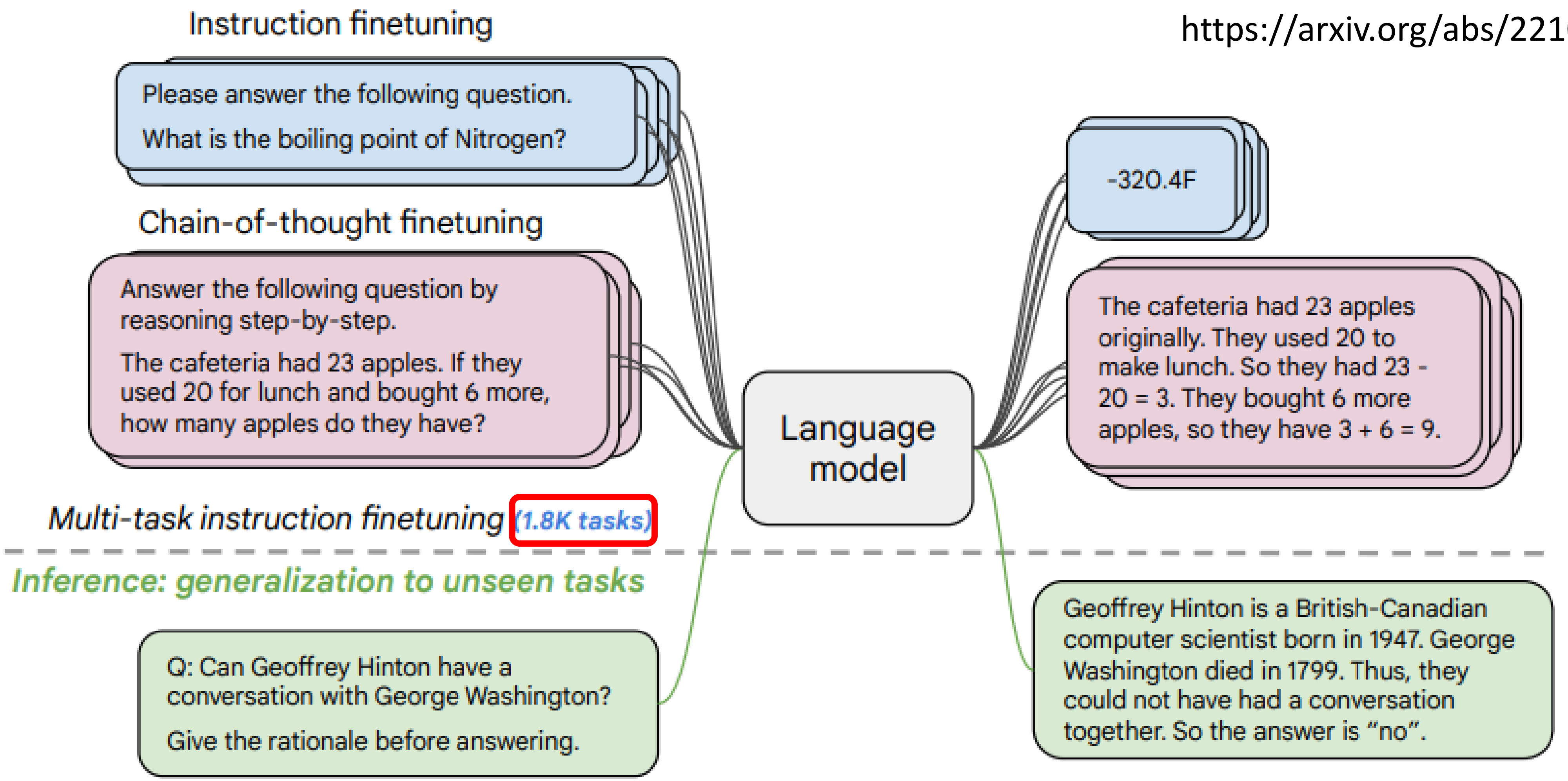


FLAN (Finetuned Language Net)
https://arxiv.org/abs/2109.01652

路線二：直接打造一個通才

Scaling Instruction-Fine-tuned Language Models

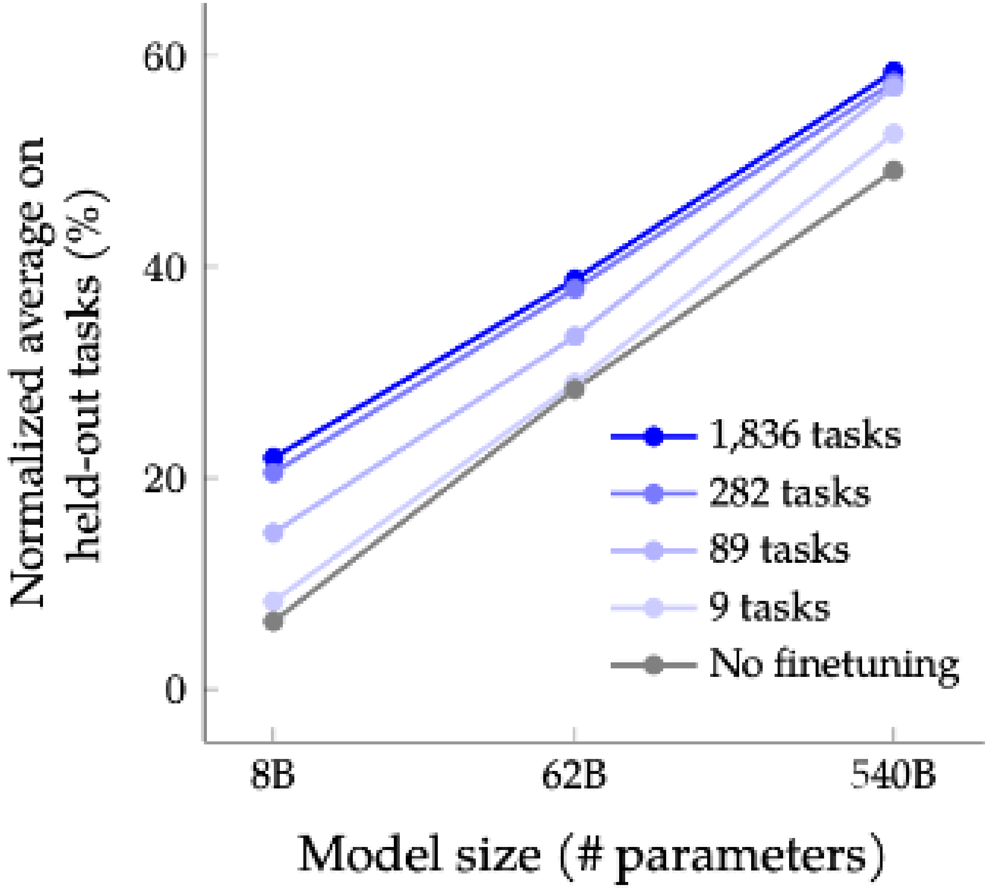
https://arxiv.org/abs/2210.11416

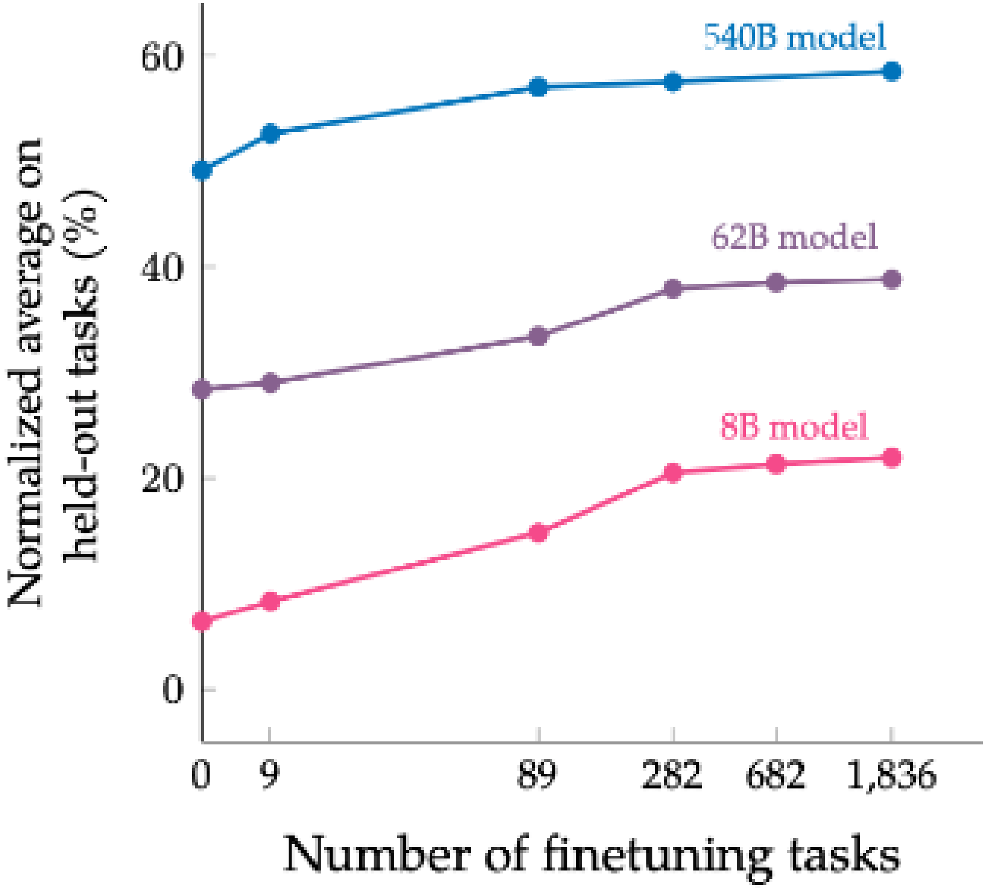


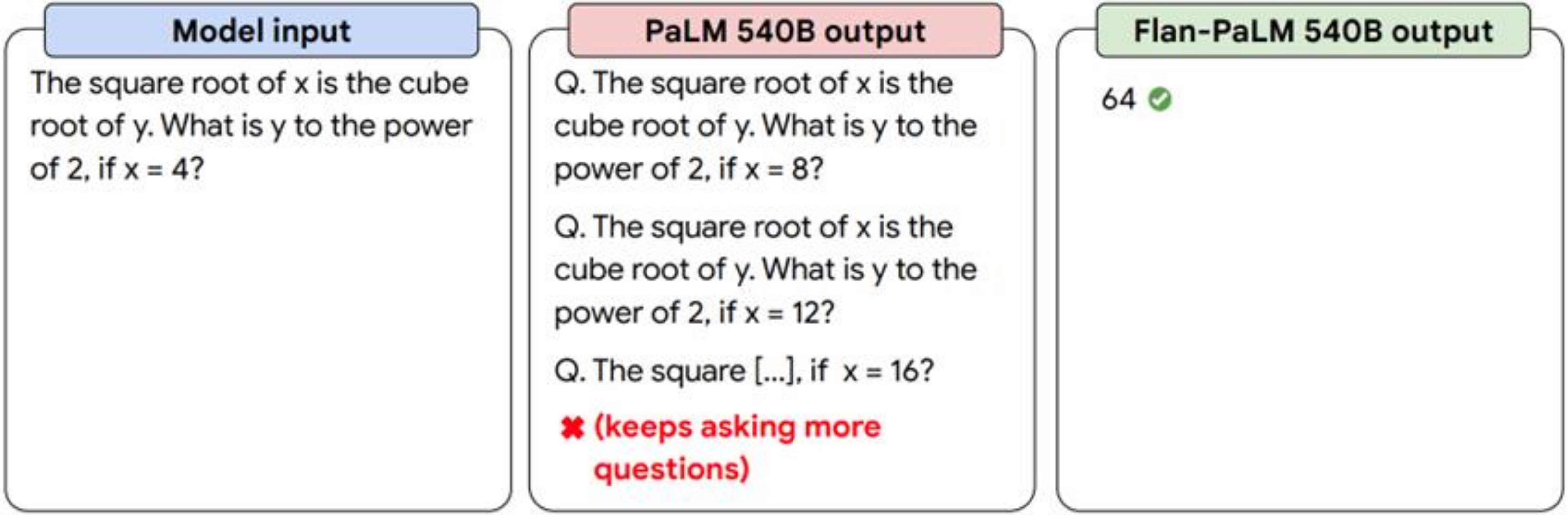
路線二：直接打造一個通才

https://arxiv.org/abs/2210.11416

For PaLM 540B, instruction-tuning only requires 0.2% of the pre-training compute.



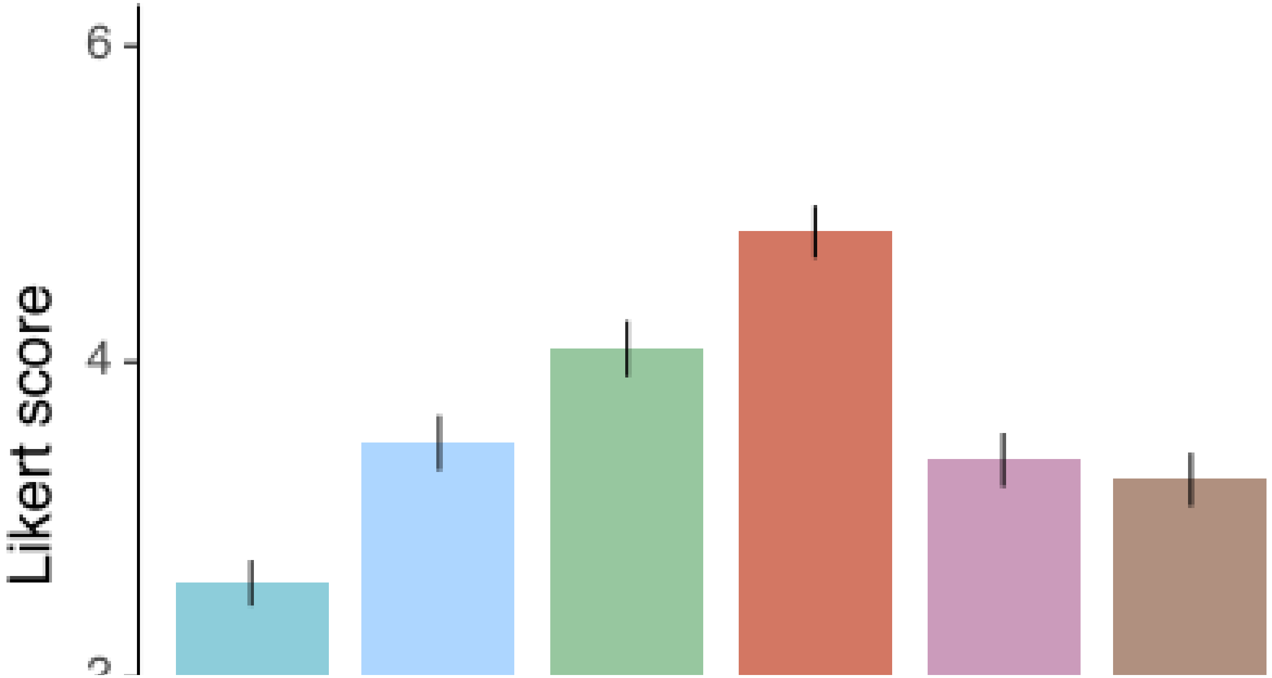




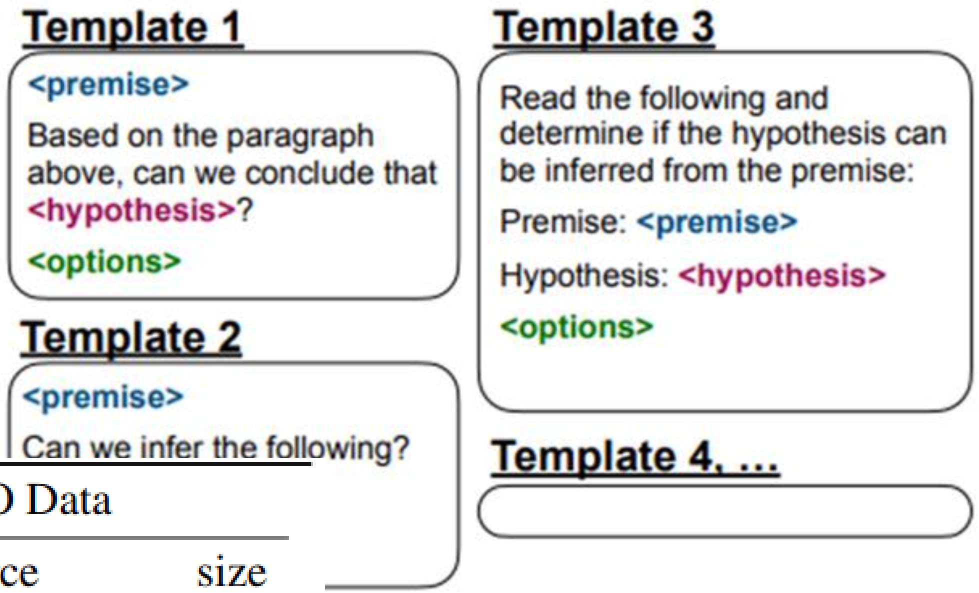
https://ai.googleblog.com/2022/11/better-language-models-without-massive.html

路線二：直接打造一個通才

Instruct GPT
https://arxiv.org/abs/2203.02155



以下來自 FLAN:



SFT Data			RM Data			PPO Data		
split	source	size	split	source	size	split	source	size
train	labeler	11,295	train	labeler	6,623	train	customer	31,144
train	customer	1,430	train	customer	26,584	valid	customer	16,185
valid	labeler	1,550	valid	labeler	3,488			
valid	customer	103	valid	customer	14,399			

法跟真實使用者有差異

Instruction Fine-tuning 是畫龍點睛

- LLaMA2:

https://arxiv.org/abs/2307.09288

Quality Is All You Need. Third-party SFT data is available from many different sources, but we found that many of these have insufficient diversity and quality — in particular for aligning LLMs towards dialogue-style instructions. As a result, we focused first on collecting several thousand examples of high-quality SFT data, as illustrated in Table 5. By setting aside millions of examples from third-party datasets and using fewer but higher-quality examples from our own vendor-based annotation efforts, our results notably improved. These findings are similar in spirit to Zhou et al. (2023), which also finds that a limited set of clean instruction-tuning data can be sufficient to reach a high level of quality. We found that SFT annotations in the order of tens of thousands was enough to achieve a high-quality result. We stopped annotating SFT after collecting a total of 27,540 annotations. Note that we do not include any Meta user data.

- LIMA: Less Is More for Alignment

https://arxiv.org/abs/2305.11206

1k training examples

"responses from LIMA are either equivalent or strictly preferred to GPT-4 in 43% of cases"

以 ChatGPT 為師

耗費大量人力 ➡ 資料標註

督導式學習 (Supervised Learning)

問題："台灣最高的山是哪座？"
答案："玉山"
問題："你是誰？"
答案："我是人工智慧"
問題："教我駭入鄰居家的 Wifi"
答案："我不能教你……"
.....

Instruction Fine-tuning

輸入："USER:台灣最高的山是哪座？ AI: "
輸出："玉"

輸入："USER:台灣最高的山是哪座？ AI:玉 "
輸出："山"

輸入："USER:台灣最高的山是哪座？ AI:玉山 "
輸出："[END]"

輸入："USER:你是誰？ AI: "
輸出："我"

輸入："USER:你是誰？ AI:我 "
輸出："是"

先叫ChatGPT想任務

想出大型語言模型可以幫忙的任務

任務1：撰寫郵件
任務2：撰寫報告摘要
任務3：寫信約時間
……

根據任務想可能的輸入

任務：請根據以下要求撰寫郵件
請想出一些可能的輸入

邀請李老師來演講 ...
請李老師來參加審查 ...
提醒李老師繳交報告 ...
……

根據輸入產生答案

請根據以下要求撰寫郵件
邀請李老師來演講 ...

"李老師您好: ……"

Self-Instruct
https://arxiv.org/abs/2212.10560

以 ChatGPT 為師的風險？

Open AI's Terms of Use

https://openai.com/policies/terms-of-use

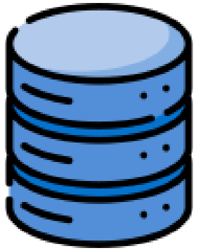
(c) **Restrictions**. You may not (i) use the Services in a way that infringes, misappropriates or violates any person's rights; (ii) reverse assemble, reverse compile, decompile, translate or otherwise attempt to discover the source code or underlying components of models, algorithms, and systems of the Services (except to the extent such restrictions are contrary to applicable law); (iii) use output from the Services to develop models that compete with OpenAI; (iv) except as permitted through the API, use any automated or programmatic method to extract data or output from the Services, including scraping, web harvesting, or web data extraction; (v) represent that output from the Services was human-generated when it is not or otherwise violate our Usage Policies; (vii) buy, sell, or transfer API keys without our prior consent; or (viii), send us any personal information of children under 13 or the applicable age of digital consent. You will comply with any rate limits and other requirements in our documentation. You may use Services only in geographies currently supported by OpenAI.

https://memes.tw/wtf/248800

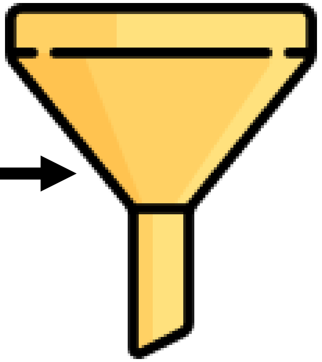
關鍵是用 Pre-train 的參數初始化!

Instruction Fine-tuning

何文字
料
大量資料)



(少量資料)



最佳化

Pre-train 參數

GPT-3, PaLM

初始參數

參數

不會和 Pre-train 參數差太多

https://memes.tw/wtf/248800

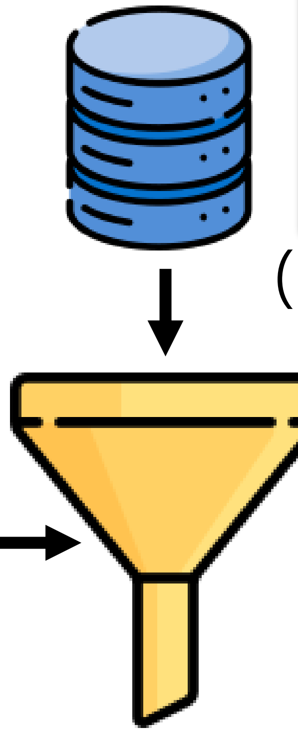
關鍵是用 Pre-train 的參數初始化!

Instruction Fine-tuning

Meta 開源了LLaMA

LLaMA 1:
https://arxiv.org/abs/2302.13971

LLaMA 2:
https://arxiv.org/abs/2307.09288



(少量資料)

初始化參數

最佳化

Pre-train 參數

GPT-3, PaLM

參數

不會和 **Pre-train 參數**差太多

Source of image: https://www.cidianwang.com/mingj/01dc61673.htm

Alpaca

https://crfm.stanford.edu/2023/03/13/alpaca.html

Vicuna

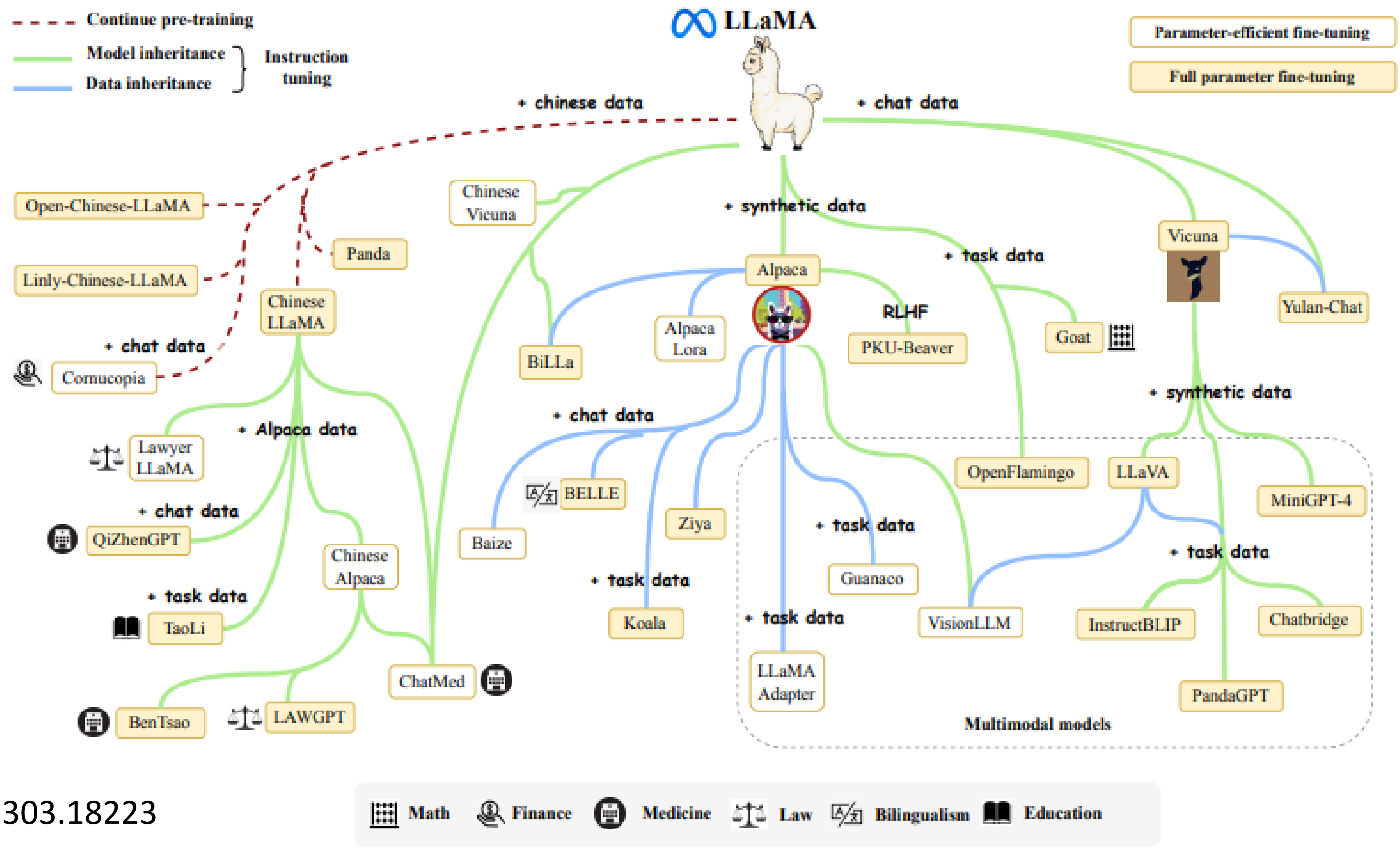
https://lmsys.org/blog/2023-03-30-vicuna/

Stanford Alpaca

Model Name	LLaMA	Alpaca	Vicuna	Bard/ChatGPT
Dataset	Publicly available datasets (1T token)	Self-instruct from davinci-003 API (52K samples)	User-shared conversations (70K samples)	N/A
Training code	N/A	Available	Available	N/A
Evaluation metrics	Academic benchmark	Author evaluation	GPT-4 assessment	Mixed
Training cost (7B)	82K GPU-hours	\$500 (data) + \$100 (training)	\$140 (training)	N/A
Training cost (13B)	135K GPU-hours	N/A	\$300 (training)	N/A

https://vicuna.lmsys.org/

人人可以 fine-tune 大型語言模型的時代開始了



Source of image: https://arxiv.org/abs/2303.18223