

# 對於大型語言模型的 兩種不同期待

**Finetune vs. Prompt**

# 語言模型

GPT



文字接龍

how are \_\_\_\_\_

BERT



文字填空

how are you

# 如何使用大型 語言模型

“A mouse riding on the head of an elephant, using reins to steer the giant creature.” (powered by Midjourney )



使用者

大型  
語言模型

# 對於大型語言模型的兩種不同期待

- 期待一：成為專才，解某一個特定任務

這堂課我們要講如何駕馭  
大型語言模型 .....



翻譯



This course  
is about .....

這堂課我們要講如何駕馭  
大型語言模型 .....



摘要

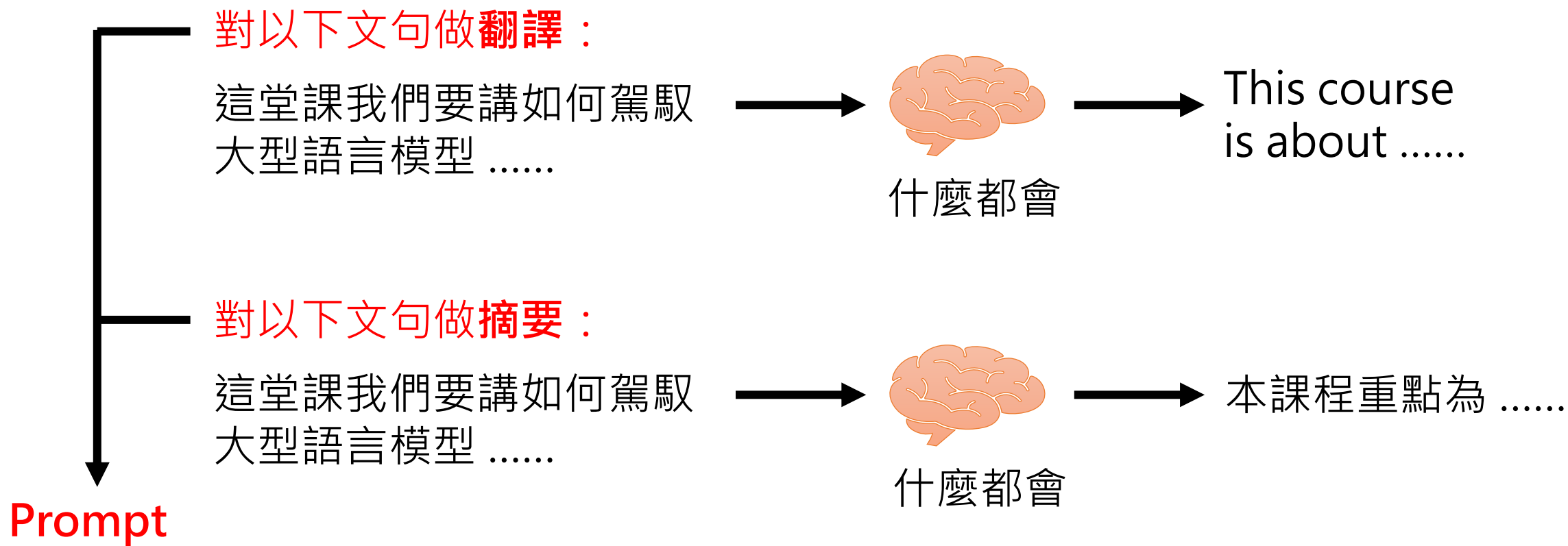


本課程重點為 .....

# 對於大型語言模型的兩種不同期待

- 期待二：成為通才

這就是今日 ChatGPT 走的路線



# 對於大型語言模型的兩種不同期待

- 期待二：成為通才

The Natural Language Decathlon: Multitask Learning as Question Answering

<https://arxiv.org/abs/1806.08730>

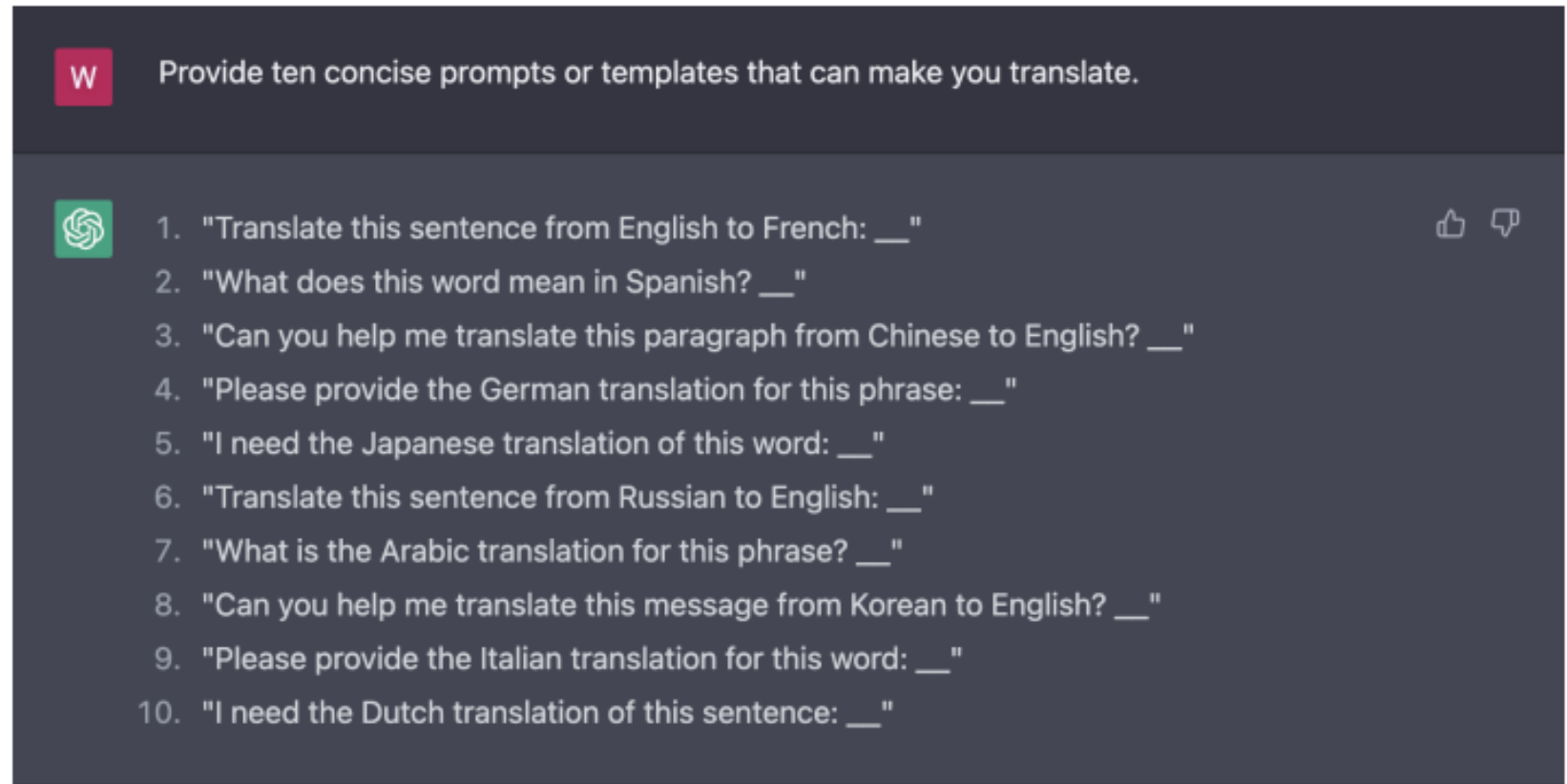
<u>Question</u>	<u>Context</u>	<u>Answer</u>	<u>Question</u>	<u>Context</u>	<u>Answer</u>
What is a major importance of Southern California in relation to California and the US?	...Southern California is a <b>major economic center</b> for the state of California and the US...	<b>major economic center</b>	What has something experienced?	Areas of the Baltic that have experienced <b>eutrophication</b> .	<b>eutrophication</b>
What is the translation from English to German?	Most of the planet is ocean water.	<b>Der Großteil der Erde ist Meerwasser</b>	Who is the illustrator of Cycle of the Werewolf?	Cycle of the Werewolf is a short novel by Stephen King, featuring illustrations by comic book artist <b>Bernie Wrightson</b> .	<b>Bernie Wrightson</b>
What is the summary?	<b>Harry Potter star Daniel Radcliffe</b> gains access to a reported <b>£320 million fortune</b> ...	<b>Harry Potter star Daniel Radcliffe gets £320M fortune...</b>	What is the change in dialogue state?	Are there any Eritrean restaurants in town?	<b>food: Eritrean</b>
Hypothesis: Product and geography are what make cream skimming work. <b>Entailment</b> , neutral, or contradiction?	Premise: Conceptually cream skimming has two basic dimensions – product and geography.	<b>Entailment</b>	What is the translation from English to SQL?	The <b>table</b> has column names... Tell me what the <b>notes</b> are for <b>South Australia</b>	<b>SELECT notes from table WHERE 'Current Slogan' = 'South Australia'</b>
Is this sentence <b>positive</b> or negative?	A stirring, funny and finally transporting re-imagining of Beauty and the Beast and 1930s horror film.	<b>positive</b>	Who had given help? <b>Susan</b> or Joan?	Joan made sure to thank Susan for all the help she had given.	<b>Susan</b>

Ask Me Anything: Dynamic Memory Networks for Natural Language Processing

<https://arxiv.org/abs/1506.07285>

# 對於大型語言模型的兩種不同期待

- 成為專才的好處：專才在單一任務上有機會贏過通才



Is ChatGPT A Good Translator?  
A Preliminary Study

<https://arxiv.org/abs/2301.08745>

# 對於大型語言模型的兩種不同期待

- 成為專才的好處：專才在單一任務上有機會贏過通才

System	De-En		Ro-En		Zh-En	
	⇒	⇐	⇒	⇐	⇒	⇐
Google	45.04	41.16	50.12	46.03	31.66	43.58
DeepL	49.23(+9.3%)	41.46(+0.7%)	50.61(+0.9%)	48.39(+5.1%)	31.22(-1.3%)	44.31(+1.6%)
Tencent	n/a	n/a	n/a	n/a	29.69(-6.2%)	46.06(+5.6%)
ChatGPT	43.71(-2.9%)	38.87(-5.5%)	44.95(-10.3%)	24.85(-46.0%)	24.73(-21.8%)	38.27(-12.1%)

System	De-Zh		Ro-Zh		De-Ro	
	⇒	⇐	⇒	⇐	⇒	⇐
Google	38.71	21.68	39.05	25.59	33.31	32.27
DeepL	40.46(+4.5%)	22.82(+5.2%)	38.95(-0.2%)	25.39(-0.7%)	35.19(+5.6%)	34.27(+6.1%)
Tencent	40.66(+5.0%)	19.44(-10.3%)	n/a	n/a	n/a	n/a
ChatGPT	34.46(-10.9%)	19.80(-8.6%)	30.84(-21.0%)	19.17(-25.0%)	33.38(+0.2%)	29.89(-7.3%)

Is ChatGPT A Good Translator?  
A Preliminary Study

<https://arxiv.org/abs/2301.08745>



# 對於大型語言模型的兩種不同期待

- 成為專才的好處：專才在單一任務上有機會贏過通才

System	COMET-22	COMETkiwi	ChrF	BLEU	COMET-22	COMETkiwi	ChrF	BLEU
		DE-EN				EN-DE		
WMT-Best	<b>85.0</b>	<b>81.4</b>	<b>58.5</b>	<b>33.4</b>	<b>87.2</b>	<b>83.6</b>	<b>64.6</b>	<b>38.4</b>
text-davinci-002	73.2	73.1	46.1	23.3	82.0	79.0	56.0	28.6
text-davinci-003	84.8*	81.2*	56.8	30.9	85.6*	82.8*	60.2*	31.8*
ChatGPT	84.8*	81.1	58.3*	33.4*	84.2	81.0	59.6	30.9
		ZH-EN				EN-ZH		
WMT-Best	81.0	77.7	<b>61.1</b>	<b>33.5</b>	<b>86.7</b>	<b>82.0</b>	<b>41.1</b>	<b>44.8</b>
text-davinci-002	74.1	73.1	49.6	20.6	84.0	79.0	32.1	36.4
text-davinci-003	<b>81.6*</b>	<b>78.9*</b>	56.0*	25.0	85.8*	81.3*	34.6	38.3
ChatGPT	81.2	78.3	56.0	25.9*	84.4	78.7	36.0*	40.3*
		RU-EN				EN-RU		
WMT-Best	<b>86.0</b>	<b>81.7</b>	<b>68.9</b>	<b>45.1</b>	<b>89.5</b>	<b>84.4</b>	<b>58.3</b>	<b>32.4</b>
text-davinci-002	77.5	76	58.7	34.9	85.4	80.9	51.6	25.1
text-davinci-003	84.8*	81.1*	64.6	38.5	86.7*	82.2*	54.0*	27.5*
ChatGPT	84.8*	81.0	66.5*	41.0*	77.6	70.4	41.1	19.0
		FR-DE				DE-FR		
WMT-Best	<b>89.5</b>	<b>80.7</b>	<b>81.2</b>	<b>64.8</b>	<b>85.7</b>	79.5	<b>74.6</b>	<b>58.4</b>
text-davinci-002	66.6	67.9	45.8	25.9	64.2	67.6	44.6	24.5
text-davinci-003	84.6	77.9	65.7*	42.5*	78.5	76.1	58.9	35.6
ChatGPT	84.7*	78.5*	65.2	42.0	81.6*	<b>79.8*</b>	60.7*	37.3*

How Good Are GPT Models at  
Machine Translation? A  
Comprehensive Evaluation

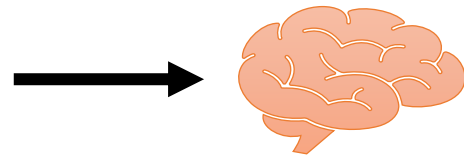
<https://arxiv.org/abs/2302.09210>

# 對於大型語言模型的兩種不同期待

- 成為通才的好處：只要重新設計 prompt 就可以快速開發新功能，不用寫程式

對以下文句做**摘要**：

這堂課我們要講如何駕馭  
大型語言模型 .....



什麼都會

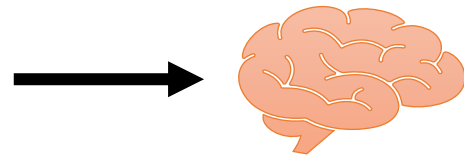
本課程重點為 .....

還是太長



請給我 **100 字**以內的**摘要**：

這堂課我們要講如何駕馭  
大型語言模型 .....



什麼都會

重點是 .....

# 兩種不同期待導致兩類不同的使用方式

- 期待一：成為專才

這堂課我們要講如何駕馭  
大型語言模型 .....



翻譯



This course  
is about .....

- 期待二：成為通才

對以下文句做**翻譯**：

這堂課我們要講如何駕馭  
大型語言模型 .....



什麼都會



This course  
is about .....

# 對預訓練模型做改造

BERT



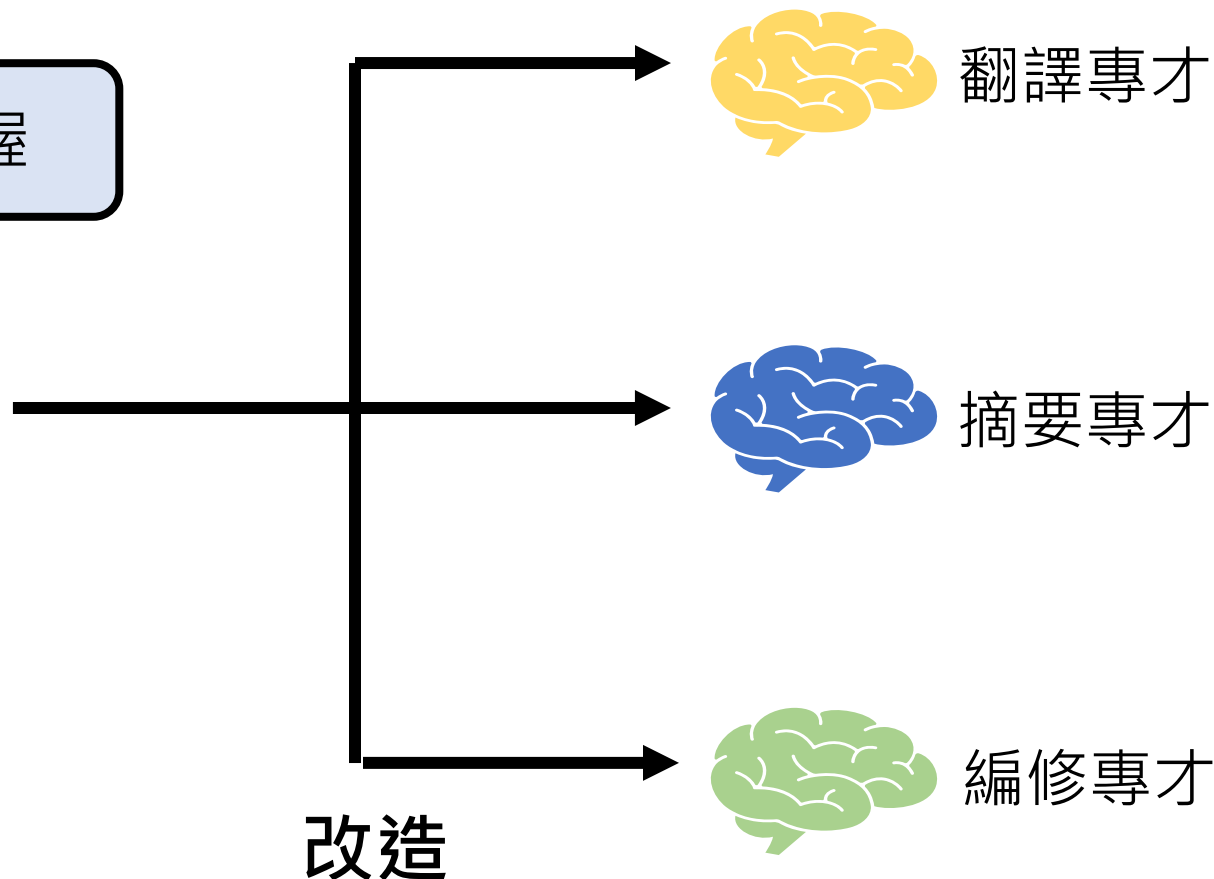
不會說一整句話喔



語言模型

文字填空

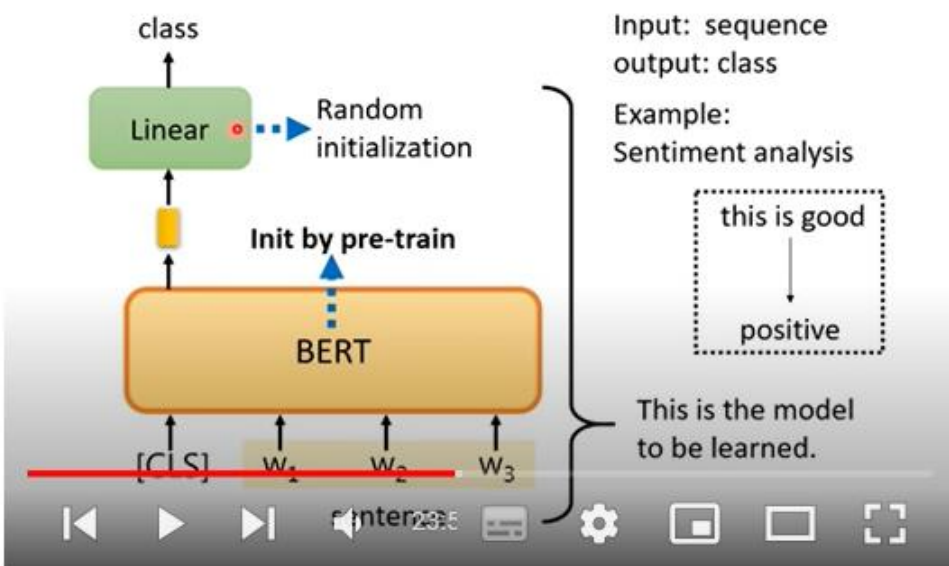
how are you



加外掛和微調參數

# 對預訓練模型做改造 — 加外掛 (Head)

## How to use BERT – Case 1

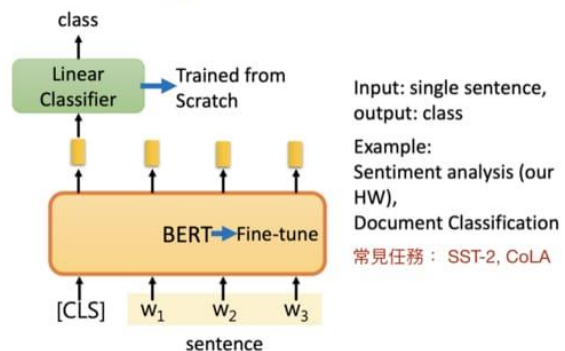


【機器學習2021】自督導式學習 (Self-supervised Learning) (二) - BERT簡介

<https://youtu.be/gh0hewYkjgo>

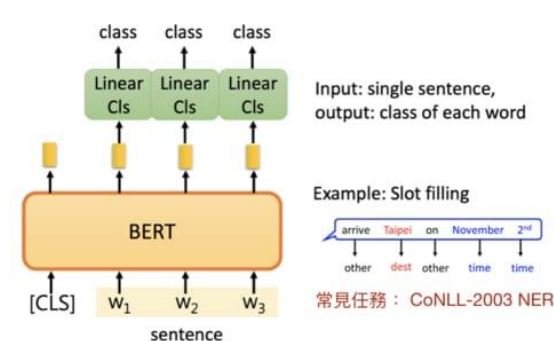
### 單一句子分類任務

bertForSequenceClassification



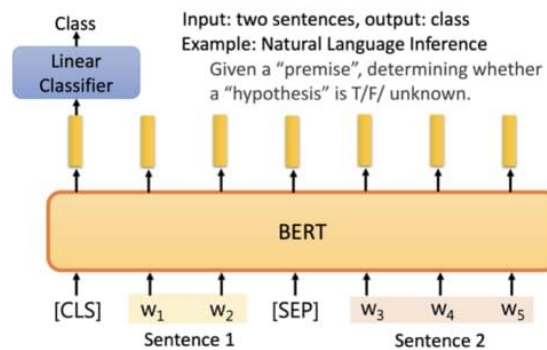
### 單一句子標註任務

bertForTokenClassification



### 成對句子分類任務

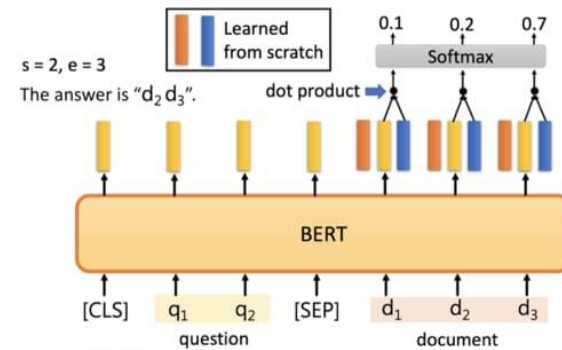
bertForSequenceClassification



常見任務: MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

### 問答任務

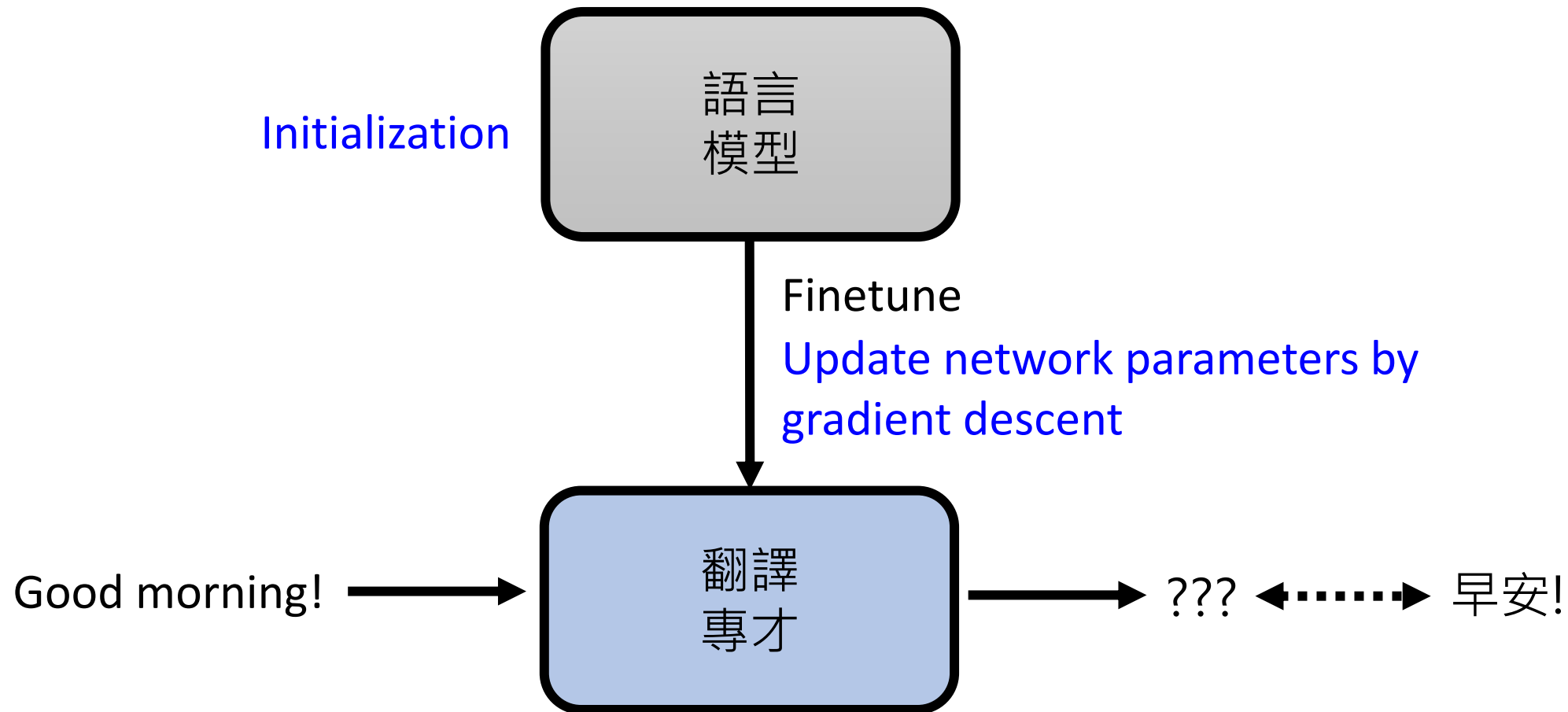
bertForQuestionAnswering



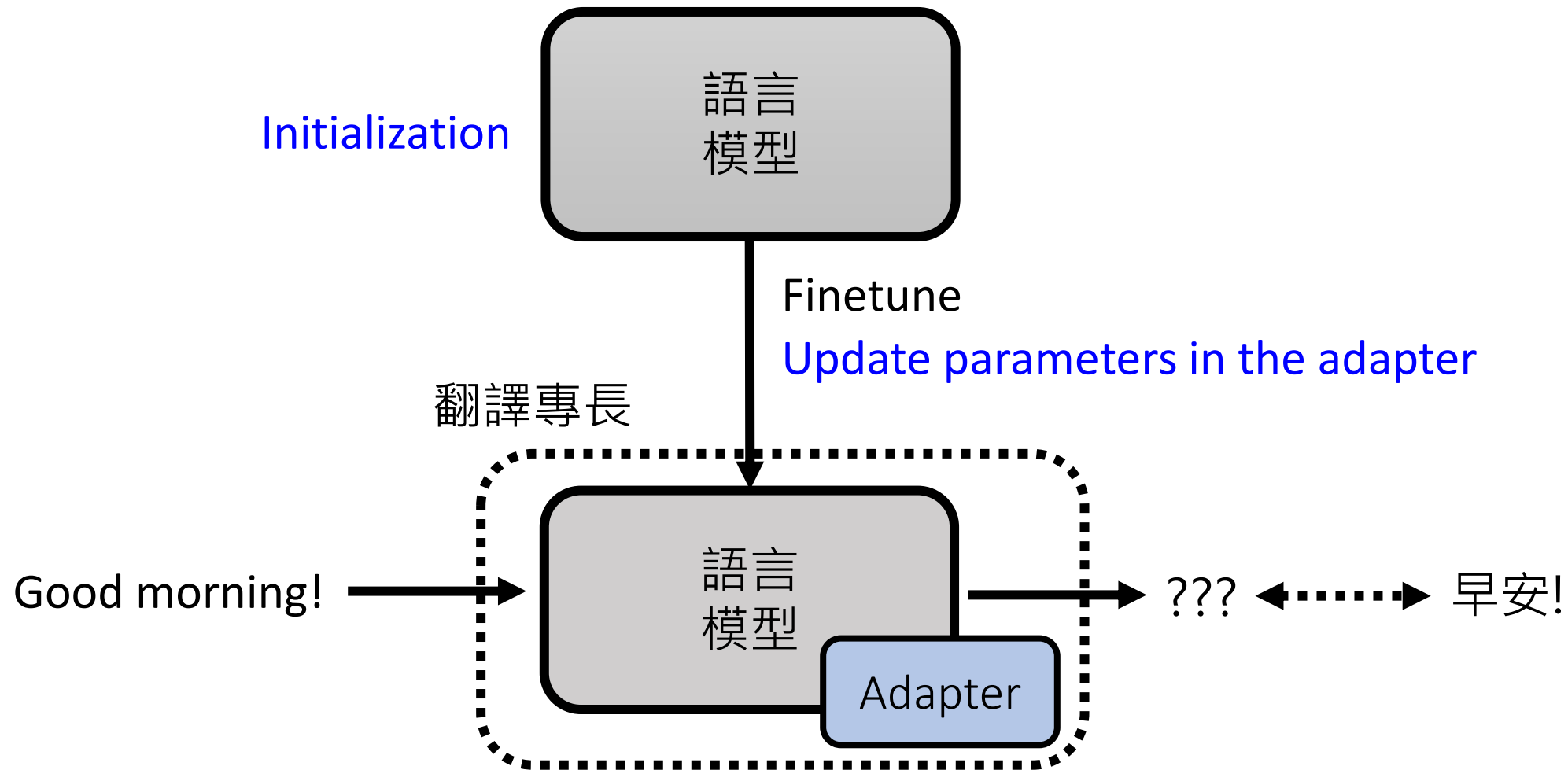
常見任務: SQuAD v1.1, SQuAD 2.0

Source of image: [https://leemeng.tw/attack\\_on\\_bert\\_transfer\\_learning\\_in\\_nlp.html](https://leemeng.tw/attack_on_bert_transfer_learning_in_nlp.html)

# 對預訓練模型做改造 — 微調 (Finetune)



# 對預訓練模型做改造 — Adapter

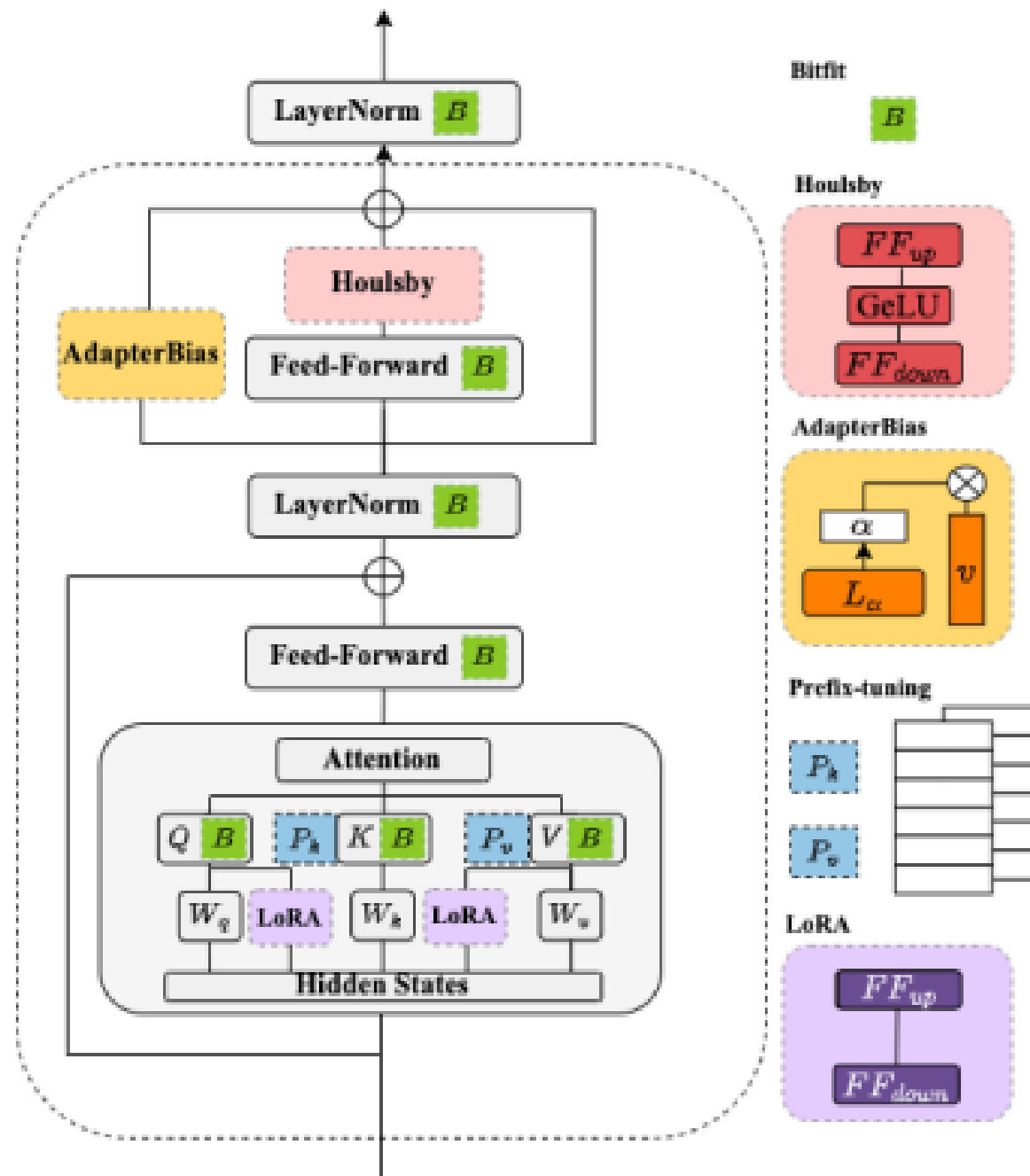


# 各種 Adapter



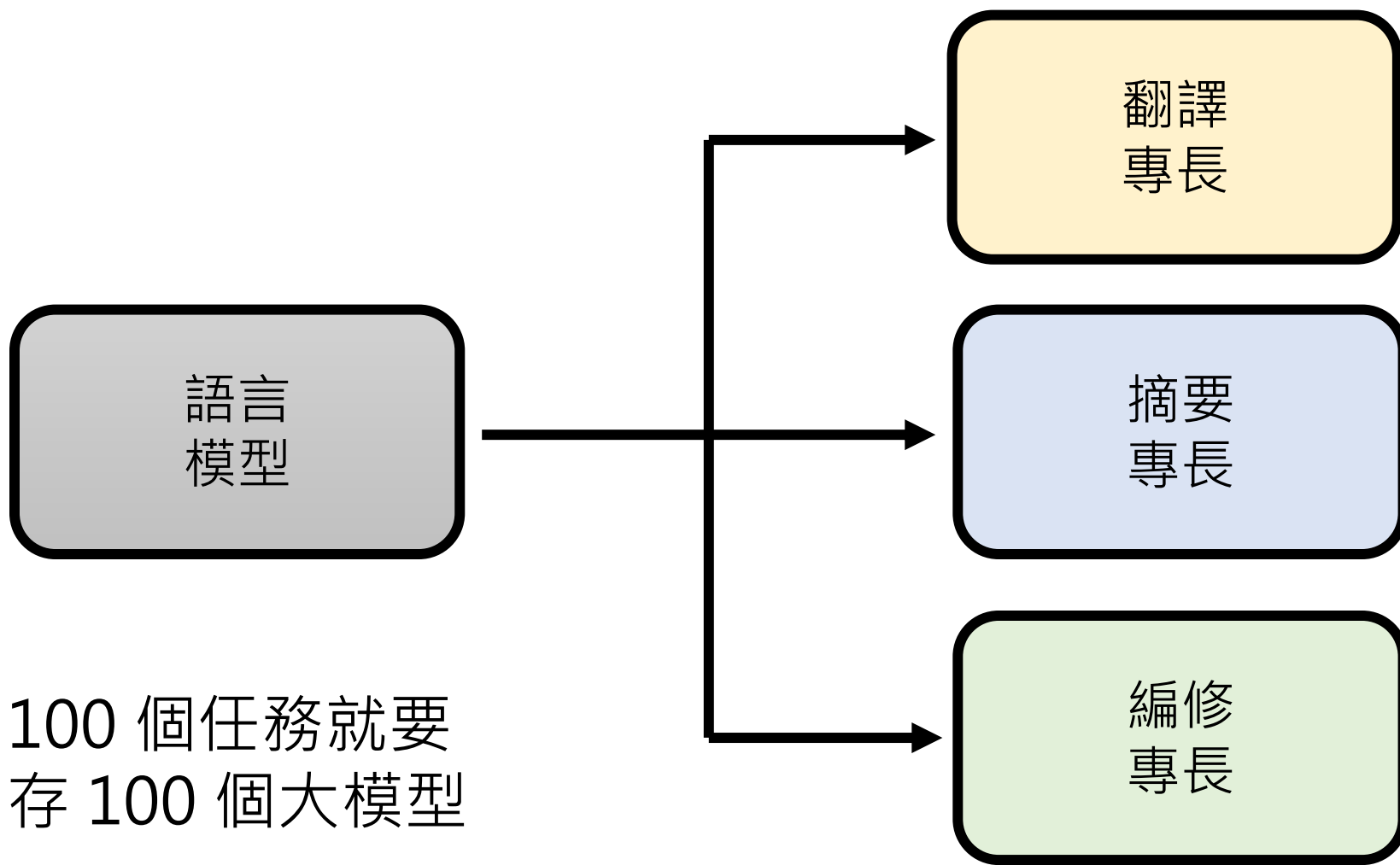
<https://adapterhub.ml/>

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

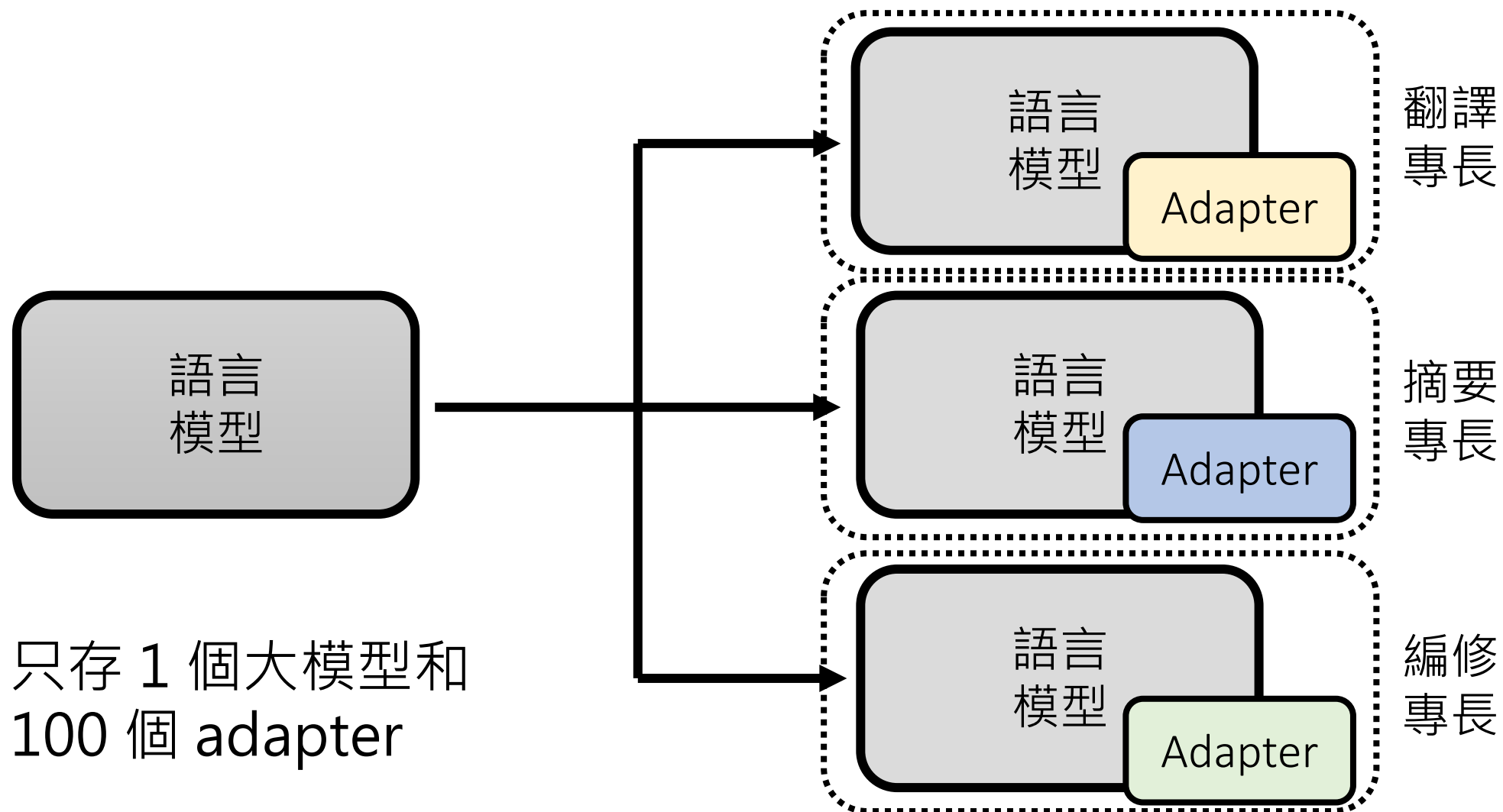




# 對預訓練模型做改造 — Adapter



# 對預訓練模型做改造 — Adapter



# 兩種不同期待導致兩類不同的使用方式

- 期待一：成為專才

這堂課我們要講如何駕馭  
大型語言模型 .....



翻譯



This course  
is about .....

- 期待二：成為通才

對以下文句做**翻譯**：

這堂課我們要講如何駕馭  
大型語言模型 .....



什麼都會



This course  
is about .....

# 機器要學會讀題目敘述或題目範例

第一部份：詞彙和結構

本部份共 15 題，每題含一個空格。請就試題冊上 A、B、C、D 四個選項中選出最適合題意的字或詞，標示在答案紙上。

例：

It's eight o'clock now. Sue \_\_\_\_\_ in her bedroom.

- A. study
- B. studies
- C. studied
- D. is studying

正確答案為 D，請在答案紙上塗黑作答。

題目敘述

Instruction Learning

範例 In-context Learning

# How to use GPT?

第一部份：詞彙和結構

本部份共 15 題，每題含一個空格。請就試題冊上 A、B、C、D 四個選項中選出最適合題意的字或詞，標示在答案紙上。

Description

例：

It's eight o'clock now. Sue \_\_\_\_\_ in her bedroom.

- A. study
- B. studies
- C. studied
- D. is studying

正確答案為 D，請在答案紙上塗黑作答。



【機器學習2021】自督導式學習 (Self-supervised Learning) (四) – GPT的野望

[https://youtu.be/WY\\_E0Sd4K80](https://youtu.be/WY_E0Sd4K80)

為什麼 GPT 系列沒有跟 BERT 一樣去微調參數？

- 一開始 OpenAI 對於 AI 就有比較高的期待？
- 因為 BERT 已經做了微調參數，只好另闢蹊徑？

# 機器要學會讀題目敘述或題目範例

第一部份：詞彙和結構

本部份共 15 題，每題含一個空格。請就試題冊上 A、B、C、D 四個選項中選出最適合題意的字或詞，標示在答案紙上。

例：

It's eight o'clock now. Sue \_\_\_\_\_ in her bedroom.

- A. study
- B. studies
- C. studied
- D. is studying

正確答案為 D，請在答案紙上塗黑作答。

題目敘述

Instruction Learning

範例 In-context Learning

# In-context Learning

今天天氣真好 分隔 正面 分隔

今天運氣真差 分隔 負面 分隔

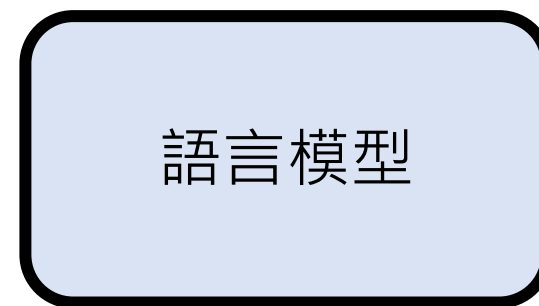
這朵花真美 分隔 正面 分隔

我真的是累了 分隔 負面 分隔

真的能從這些例子學習？

給一些例子

我感到非常高興



語言模型

正面

任務：情感分析

# In-context Learning

Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Ref: <https://arxiv.org/abs/2202.12837>

今天天氣真好 分隔 負面 分隔

今天運氣真差 分隔 正面 分隔

這朵花真美 分隔 負面 分隔

我真的是累了 分隔 正面 分隔

我感到非常高興

故意給錯誤標註

給一些例子

語言模型

正面

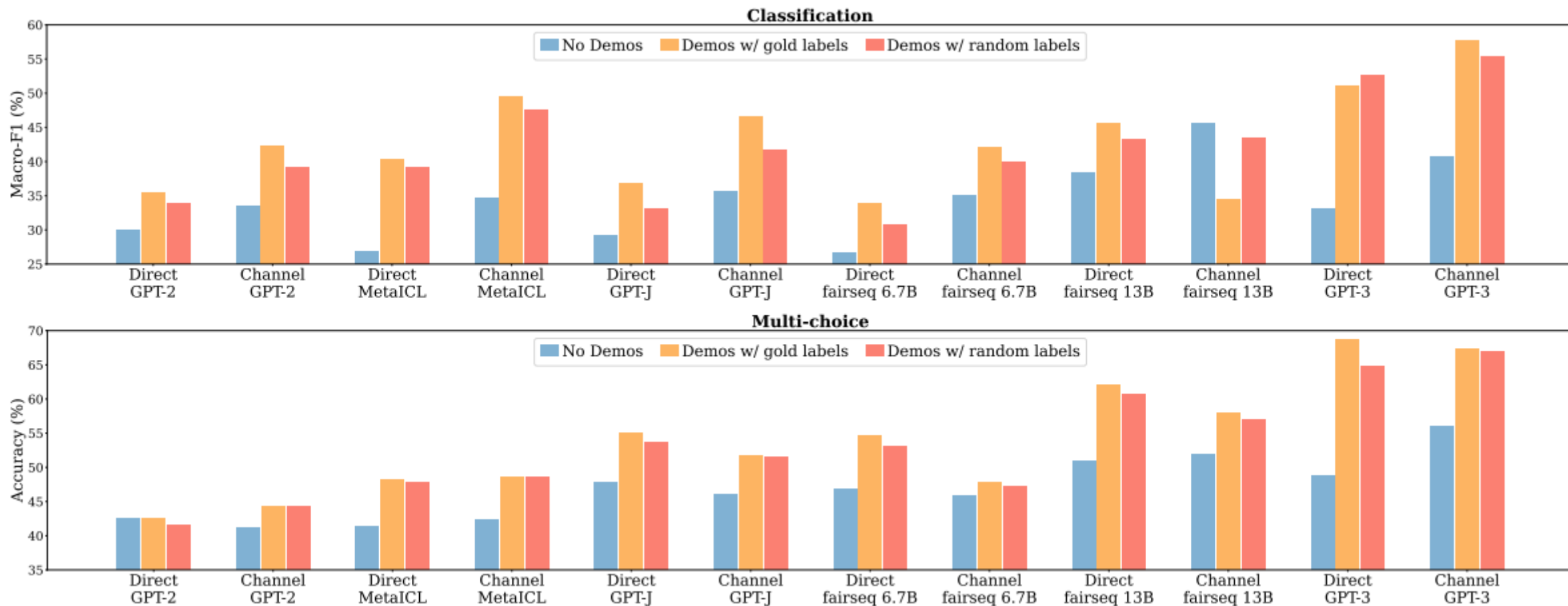
任務：情感分析



# In-context Learning

Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Ref: <https://arxiv.org/abs/2202.12837>



# In-context Learning

Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Ref: <https://arxiv.org/abs/2202.12837>

&#(&\$%@# 分隔 負面 分隔

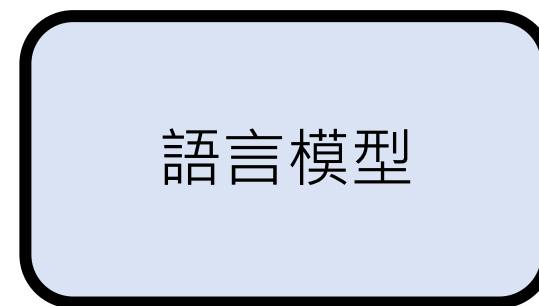
&#(&\$%@# 分隔 正面 分隔

&#(^\$@# 分隔 負面 分隔

#&(&\$%&) 分隔 正面 分隔

故意給錯誤標註  
給一些例子  
故意給無關的輸入

我感到非常高興



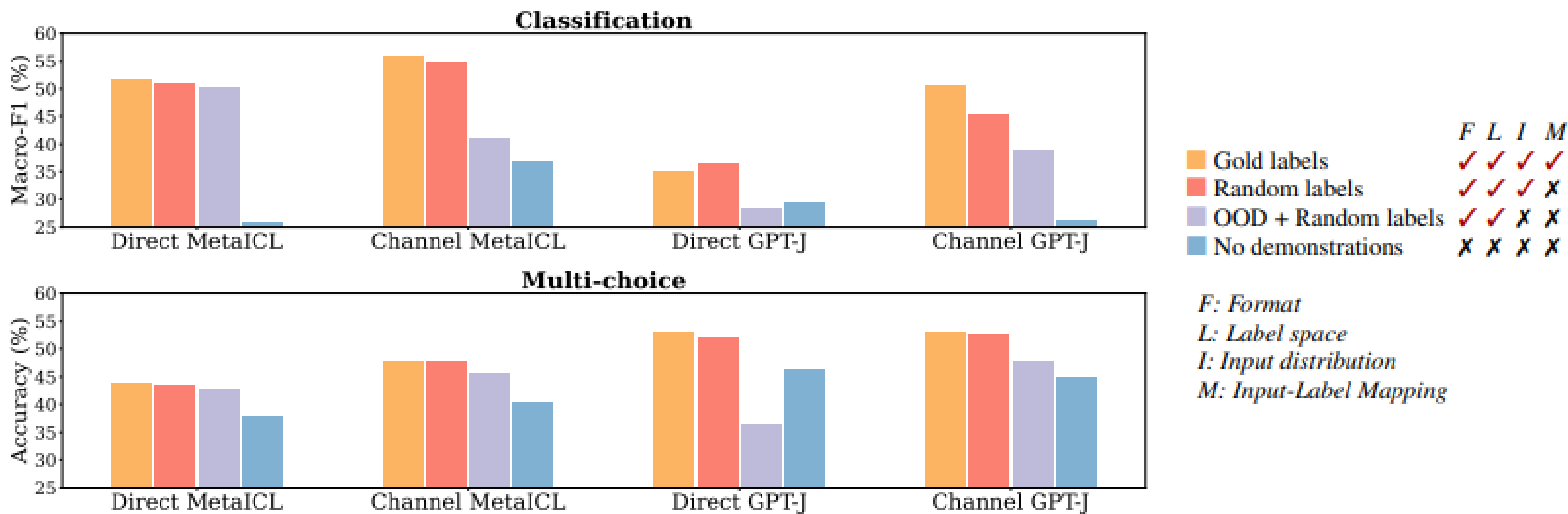
正面

任務：情感分析

# In-context Learning

Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Ref: <https://arxiv.org/abs/2202.12837>



# In-context Learning

Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Ref: <https://arxiv.org/abs/2202.12837>

今天天氣真好 分隔 正面 分隔

今天運氣真差 分隔 負面 分隔

這朵花真美 分隔 正面 分隔

我真的是累了 分隔 負面 分隔

我感到非常高興

本來就會做情感分析，只是  
需要被指出需要做情感任務

語言模型

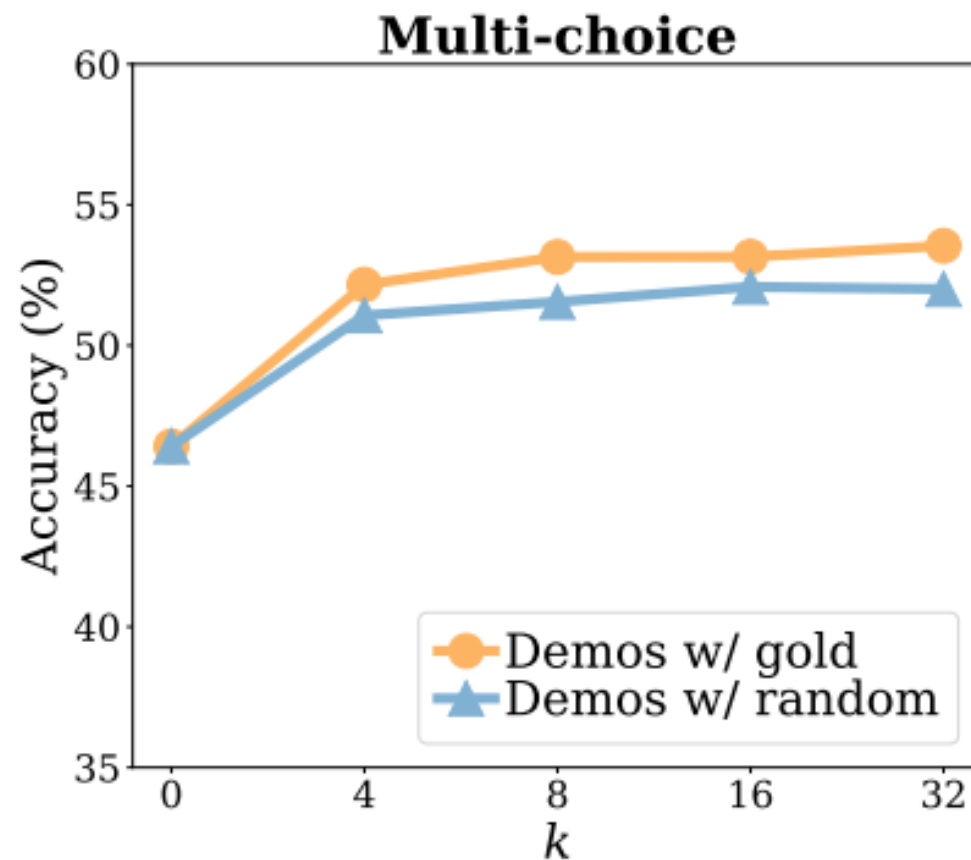
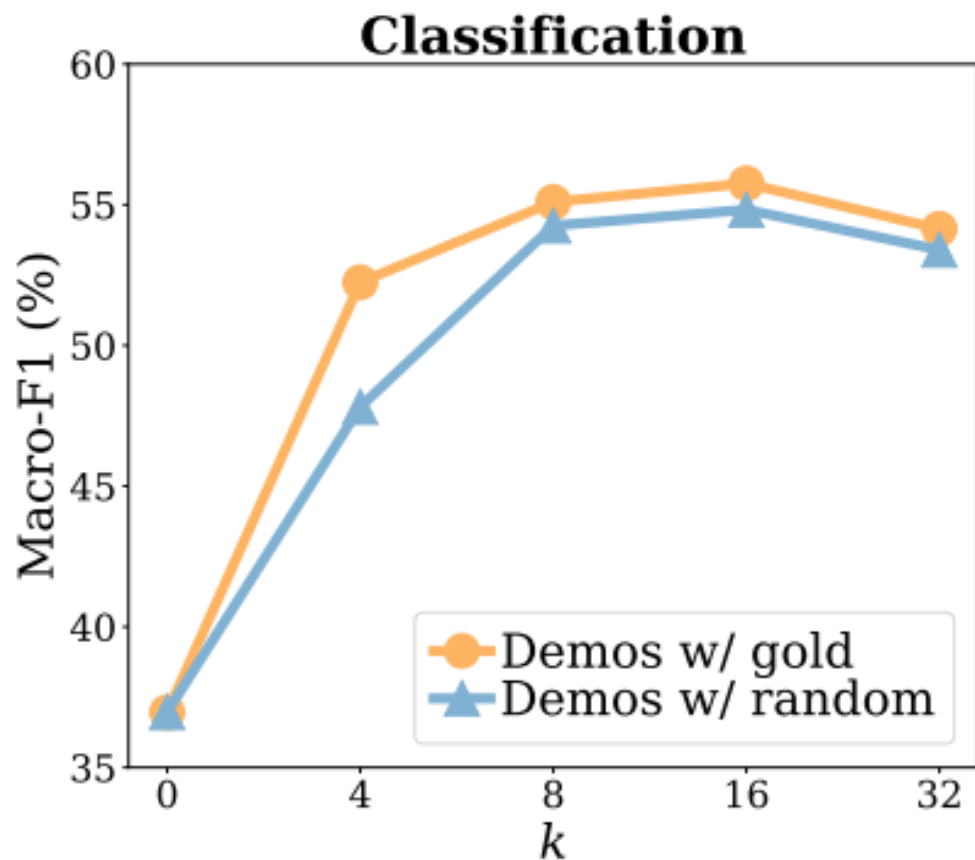
正面

任務：情感分析

# In-context Learning

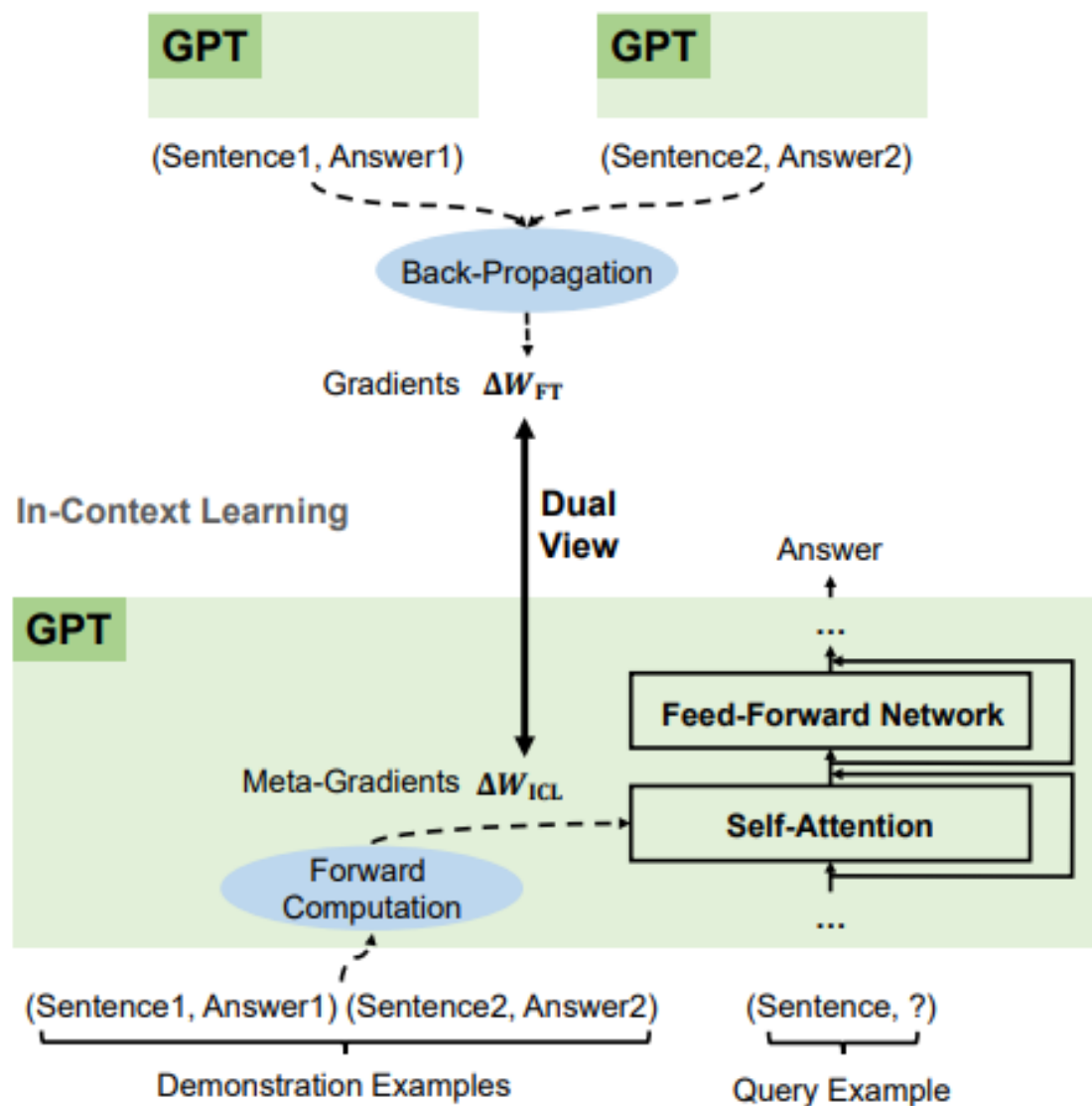
Rethinking the Role of Demonstrations: What Makes In-Context Learning Work?

Ref: <https://arxiv.org/abs/2202.12837>



# In-context Learning

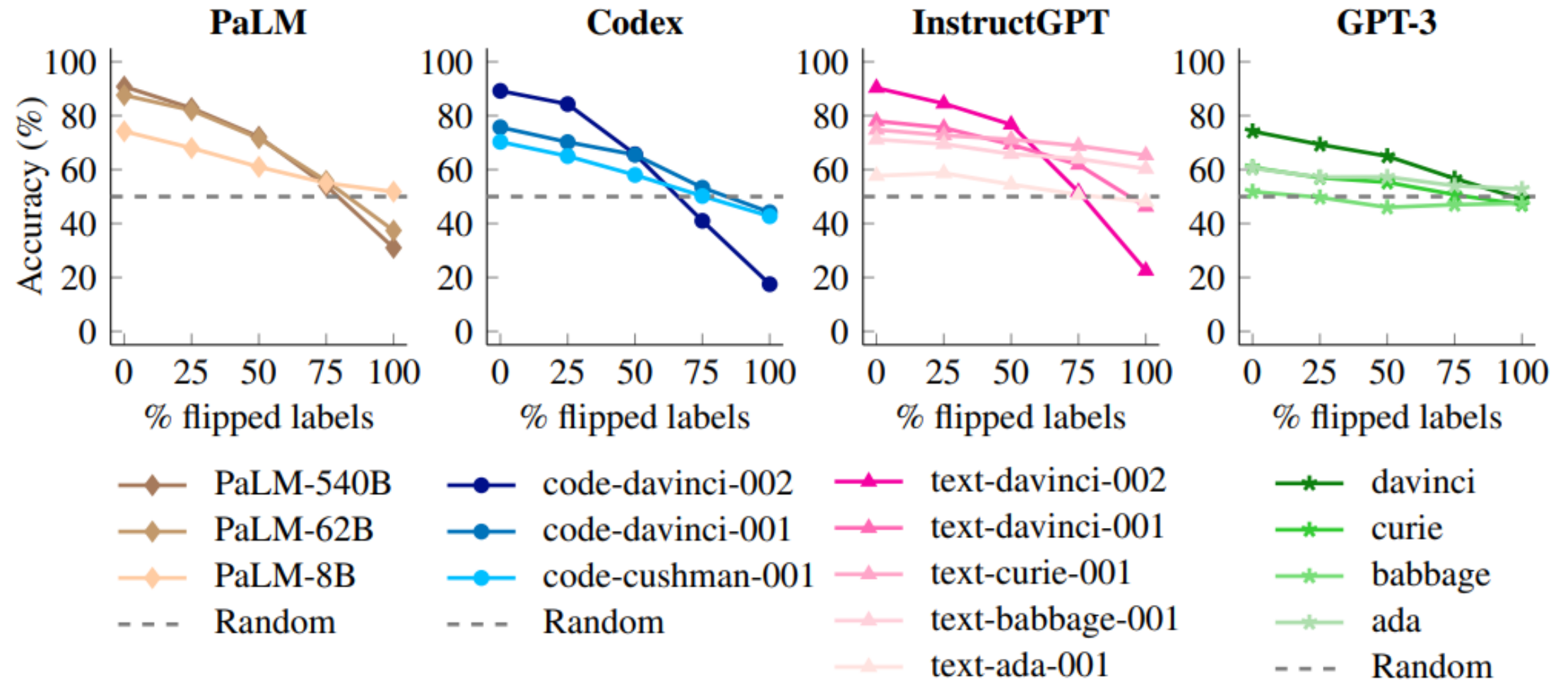
## Finetuning



<https://arxiv.org/abs/2212.10559>

<https://arxiv.org/abs/2211.15661>

# In-context Learning



# In-context Learning

- Classification!

Input: 680, 841, 842, 496, 204, 985, 546, 275, 453, 835, 644, 1, 308, 5, 65, 160

Output: Bar

Input: 193, 101, 270, 957, 670, 407, 104, 23, 569, 708, 700, 395, 481, 105, 234, 785

Output: Foo

Input: 16, 409, 28, 668, 53, 342, 813, 181, 963, 728, 558, 420, 975, 686, 395, 931

Output: Bar

Input: 448, 421, 190, 246, 413, 766, 463, 332, 935, 911, 304, 244, 876, 95, 236, 695

Output: Foo

Input: 632, 318, 49, 138, 602, 508, 924, 227, 325, 767, 108, 254, 475, 298, 202, 989

Output: Foo

Input: 412, 140, 30, 508, 837, 707, 338, 669, 835, 177, 312, 800, 526, 298, 214, 259

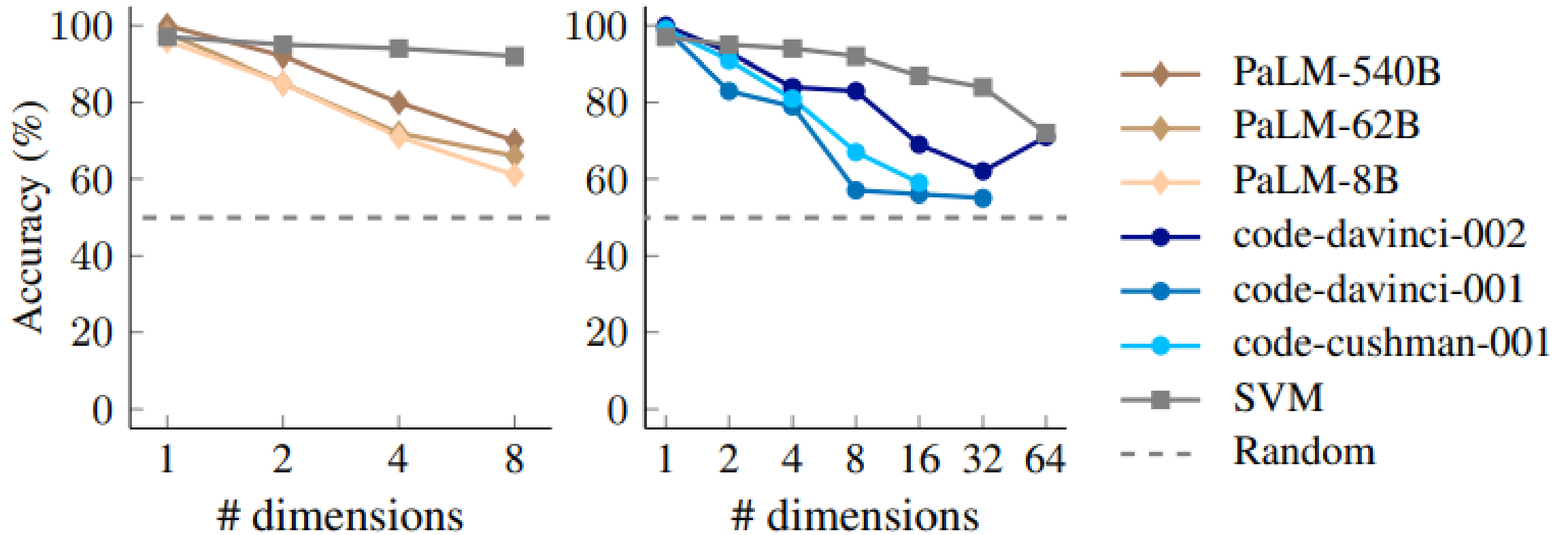
Output:

**Answer:**

Bar

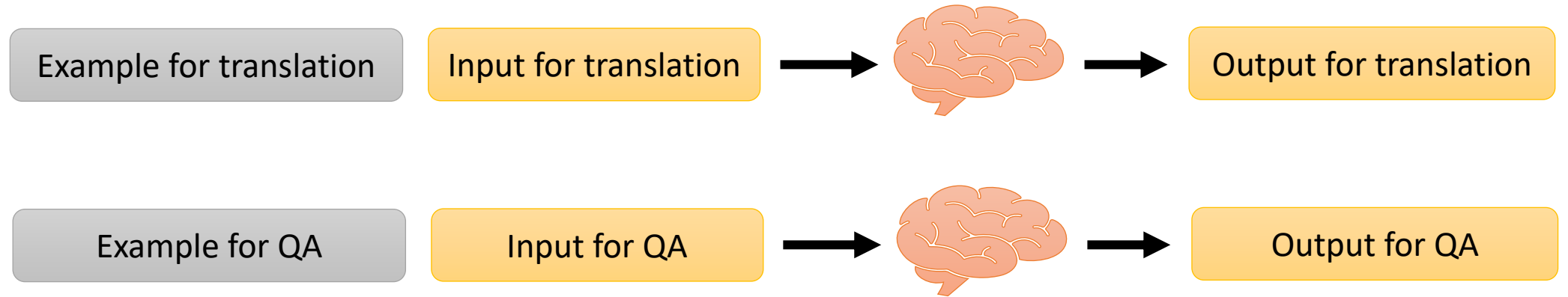


# In-context Learning

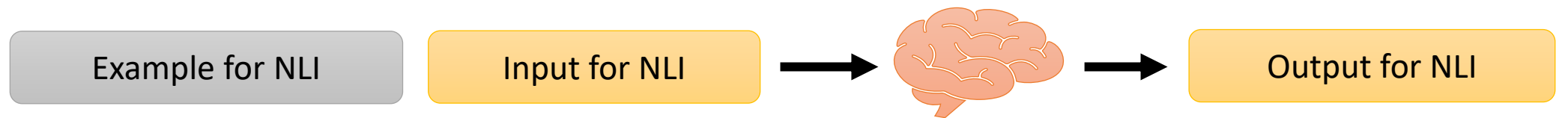


# Learning In-context Learning

**Training**



**Testing**



NLI = Natural Language Inference

# 機器要學會讀題目敘述或題目範例

第一部份：詞彙和結構

本部份共 15 題，每題含一個空格。請就試題冊上 A、B、C、D 四個選項中選出最適合題意的字或詞，標示在答案紙上。

例：

It's eight o'clock now. Sue \_\_\_\_\_ in her bedroom.

- A. study
- B. studies
- C. studied
- D. is studying

正確答案為 D，請在答案紙上塗黑作答。

題目敘述

Instruction Learning

範例 In-context Learning

# Instruction-tuning

Training

對以下文句做翻譯：這堂課我們  
要講如何駕馭大型語言模型 .....



This course  
is about .....

對以下文句做摘要：這堂課我們  
要講如何駕馭大型語言模型 .....



本課程重點  
為 .....

Testing

請幫我編修以下文句："How is  
you?"



"How are  
you?"

## Summarization

The picture appeared on the wall of a Poundland store on Whymark Avenue [...] How would you rephrase that in a few words?

## Sentiment Analysis

Review: We came here on a Saturday night and luckily it wasn't as packed as I thought it would be [...] On a scale of 1 to 5, I would give this a

## Question Answering

I know that the answer to "What team did the Panthers defeat?" is in "The Panthers finished the regular season [...]". Can you tell me what it is?

<https://arxiv.org/abs/2110.08207>

Graffiti artist Banksy is believed to be behind [...]

4

Arizona Cardinals

T<sub>0</sub>

Yes

Multi-task training

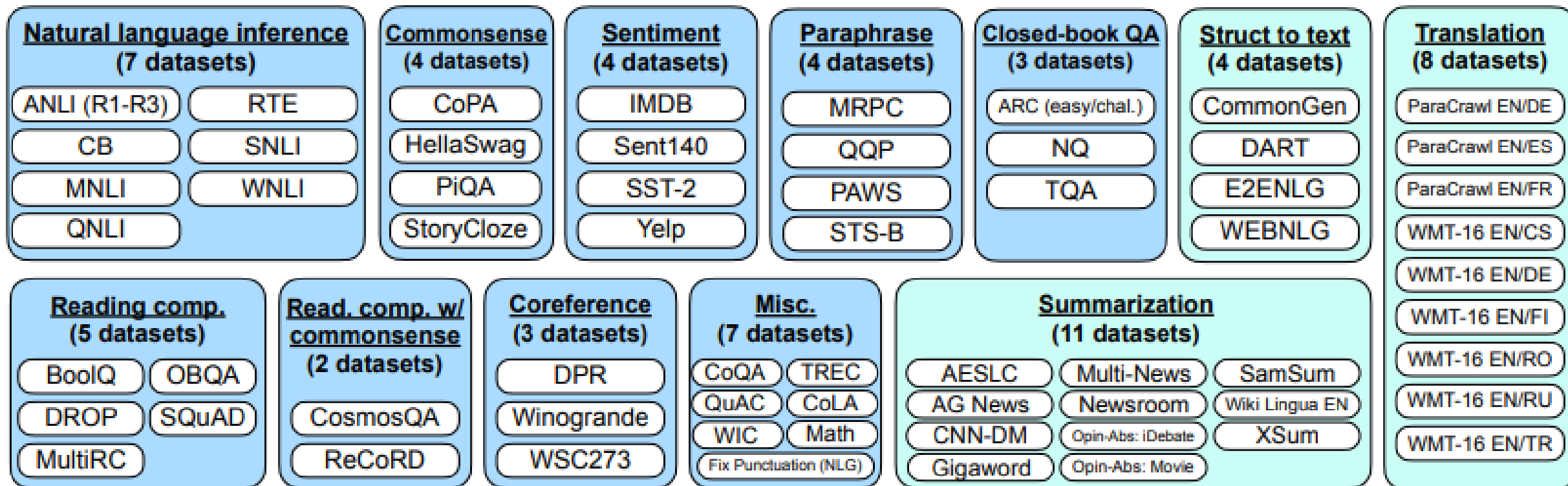
Zero-shot generalization

## Natural Language Inference

Suppose "The banker contacted the professors and the athlete". Can we infer that "The banker contacted the professors"?

# Instruction-tuning

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



# Instruction-tuning

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

## Premise

Russian cosmonaut Valery Polyakov set the record for the longest continuous amount of time spent in space, a staggering 438 days, between 1994 and 1995.

## Hypothesis

Russians hold the record for the longest stay in space.

## Target

Entailment  
Not entailment



Options:

- yes
- no



## Template 1

<premise>

Based on the paragraph above, can we conclude that <hypothesis>?

<options>

## Template 2

<premise>

Can we infer the following?

<hypothesis>

<options>

## Template 3

Read the following and determine if the hypothesis can be inferred from the premise:

Premise: <premise>

Hypothesis: <hypothesis>

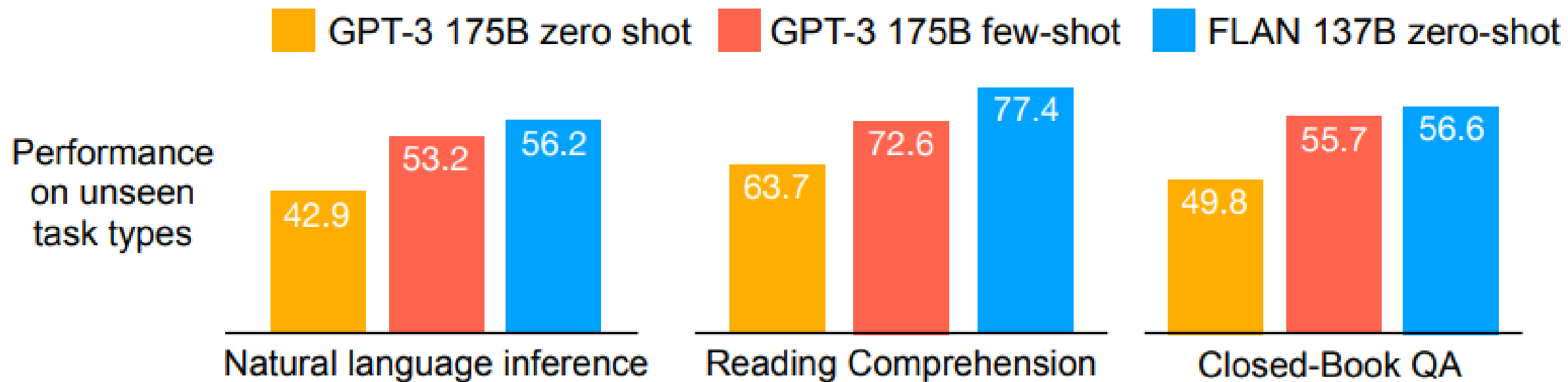
<options>

## Template 4, ...



# Instruction-tuning

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





# Chain of Thought (CoT) Prompting

## Standard Prompting

### Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

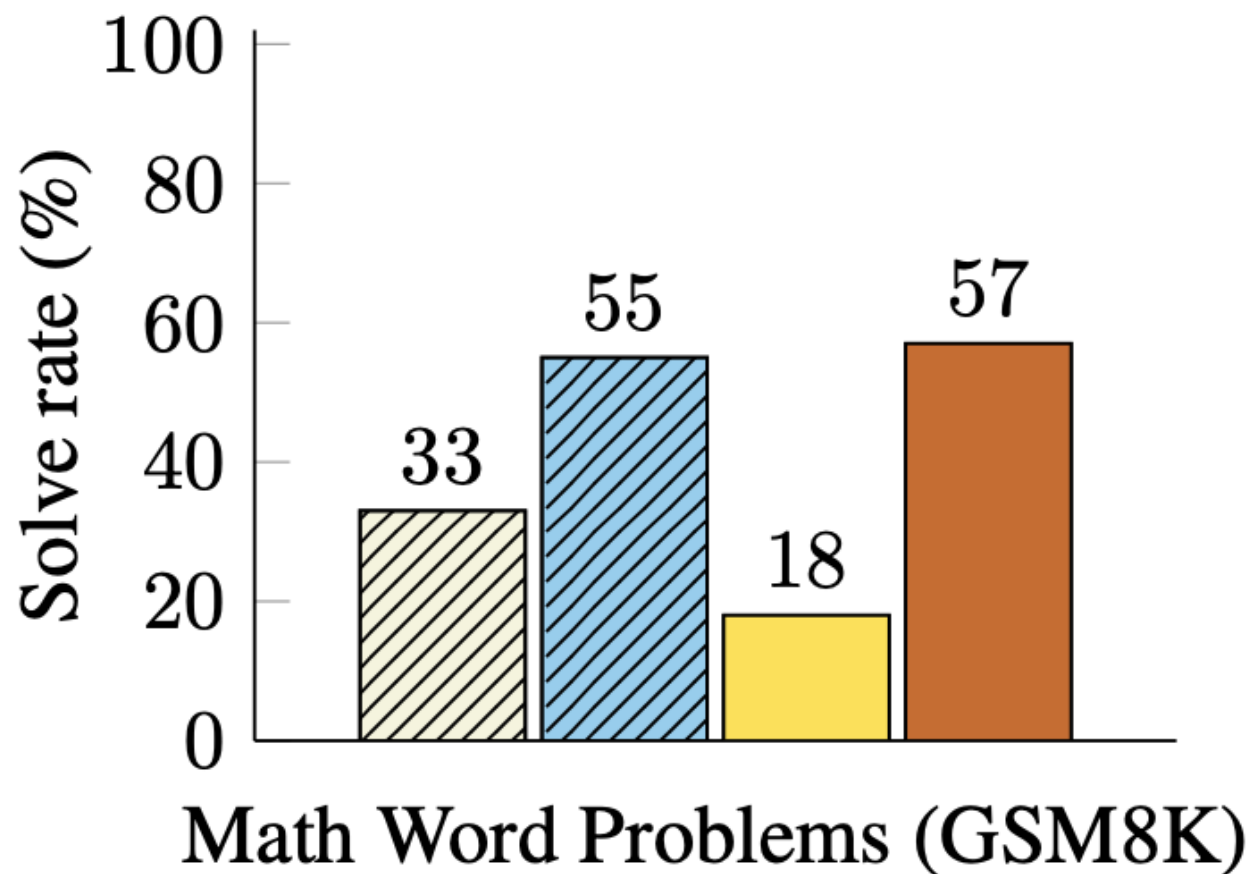
A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

### Model Output

A: The answer is 27. ❌

- Finetuned GPT-3 175B
- Prior best
- PaLM 540B: standard prompting
- PaLM 540B: chain-of-thought prompting



# Chain of Thought (CoT) Prompting

(a) Few-shot

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

---

*(Output) The answer is 8. X*

(b) Few-shot-CoT

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.

Q: A juggler can juggle 16 balls. Half of the balls are golf balls, and half of the golf balls are blue. How many blue golf balls are there?

A:

---

*(Output) The juggler can juggle 16 balls. Half of the balls are golf balls. So there are  $16 / 2 = 8$  golf balls. Half of the golf balls are blue. So there are  $8 / 2 = 4$  blue golf balls. The answer is 4. ✓*

# Chain of Thought (CoT) Prompting

## Self-consistency

**Q:** If there are 3 cars in the parking lot and 2 more cars arrive, how many cars are in the parking lot?

**A:** There are 3 cars in the parking lot already. 2 more arrive. Now there are  $3 + 2 = 5$  cars. The answer is 5.

...

**Q:** Janet's ducks lay 16 eggs per day. She eats three for breakfast every morning and bakes muffins for her friends every day with four. She sells the remainder for \$2 per egg. How much does she make every day?

**A:**

Language model

Sample a diverse set of reasoning paths

She has  $16 - 3 - 4 = 9$  eggs left. So she makes  $\$2 * 9 = \$18$  per day.

The answer is \$18.

This means she she sells the remainder for  $\$2 * (16 - 4 - 3) = \$26$  per day.

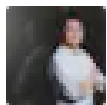
The answer is \$26.

She eats 3 for breakfast, so she has  $16 - 3 = 13$  left. Then she bakes muffins, so she has  $13 - 4 = 9$  eggs left. So she has  $9 \text{ eggs} * \$2 = \$18$ .

The answer is \$18.

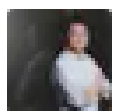
Marginalize out reasoning paths to aggregate final answers

The answer is \$18.



雞、鴨、兔共30隻，72條腿。其中雞的數量是鴨的2倍，那麼雞有幾隻？直接給答案，不列計算過程

雞 16 隻 (鴨 8 隻、兔 6 隻)



雞、鴨、兔共30隻，72條腿。其中雞的數量是鴨的2倍，那麼雞有幾隻？請詳列計算過程，再給答案



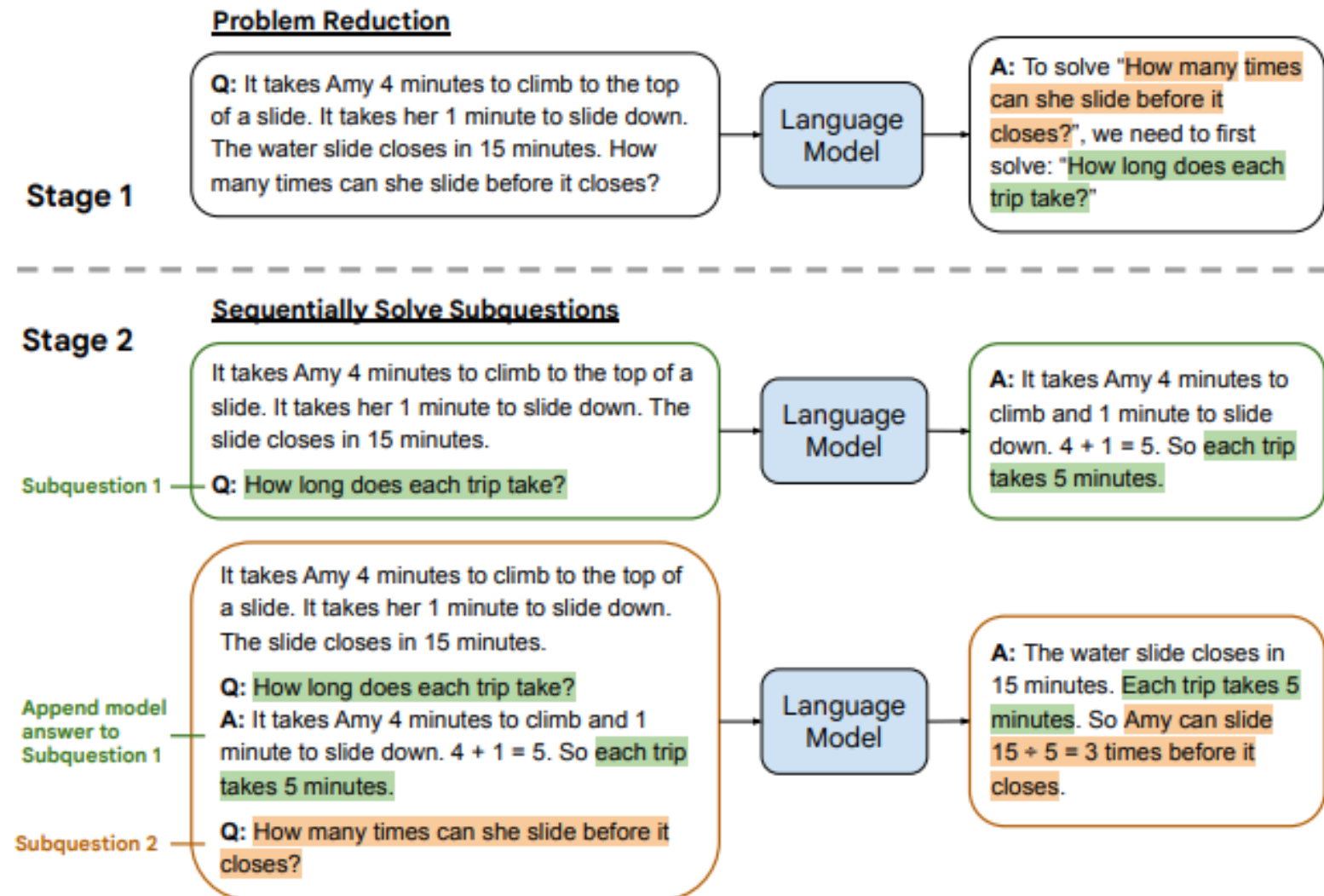
設雞的數量為 $x$ ，鴨的數量為 $y$ ，兔的數量為 $z$ 。

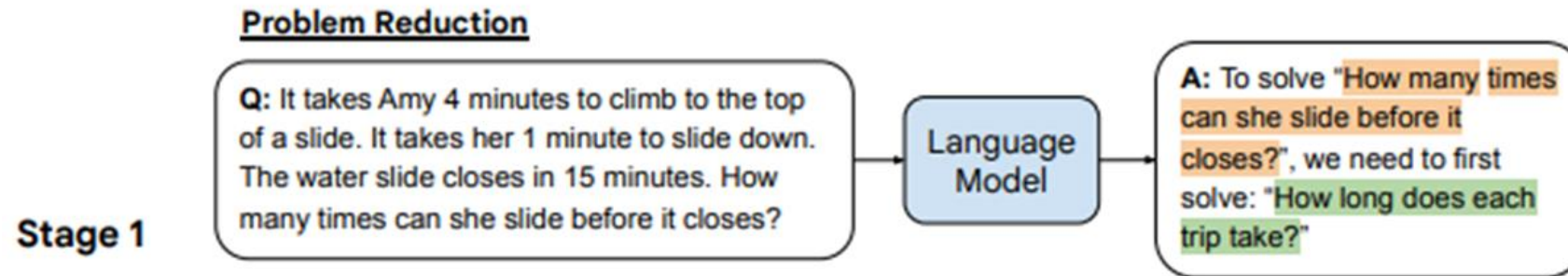
五次答案為：18, 8, 12, 7, 2

由題意可得以下三個方程：

# Chain of Thought (CoT) Prompting

## Least-to-most prompting





**Q:** The Seahawks played the San Francisco 49ers. In the first quarter, the 'Hawks RB Julius Jones got a 27-yard TD run, along with DT Craig Terrill returning a fumble 9 yards for a touchdown. In the third quarter, the 49ers almost rallied as RB T. J. Duckett made a 12-yard TD pass to Lucas Nelly, along with Mare kicking a 32-yard field goal. How many yards do the shortest touchdown run and the longest touchdown pass combine for?

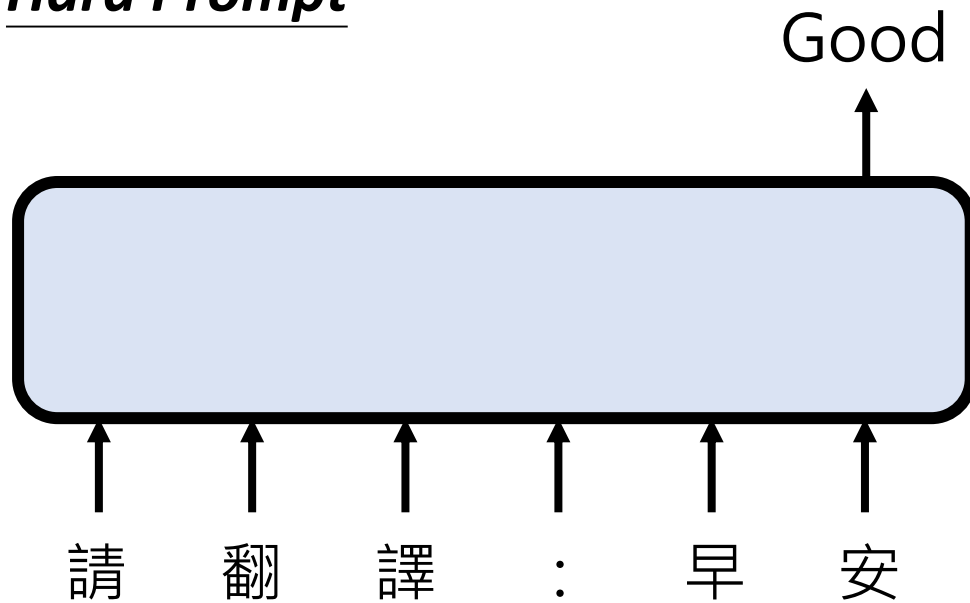
**A:** To answer the question "How many yards do the shortest touchdown run and the longest touchdown pass combine for?", we need to know: "How many yards was the shortest touchdown run?", "How many yards was the longest touchdown pass?"

**Q:** The Steelers went home for an AFC North duel with the Baltimore Ravens. Pittsburgh would deliver the opening punch in the first quarter with a 1-yard touchdown run from running back Rashard Mendenhall. The Ravens would make it even in the second quarter as running back Willis McGahee got a 9-yard touchdown run. The Ravens kicker Billy Cundiff got a 45-yard field goal in the second quarter and a 33-yard field goal in the third quarter. Game ended with a scoreless fourth quarter. How many points did the Ravens have at halftime?

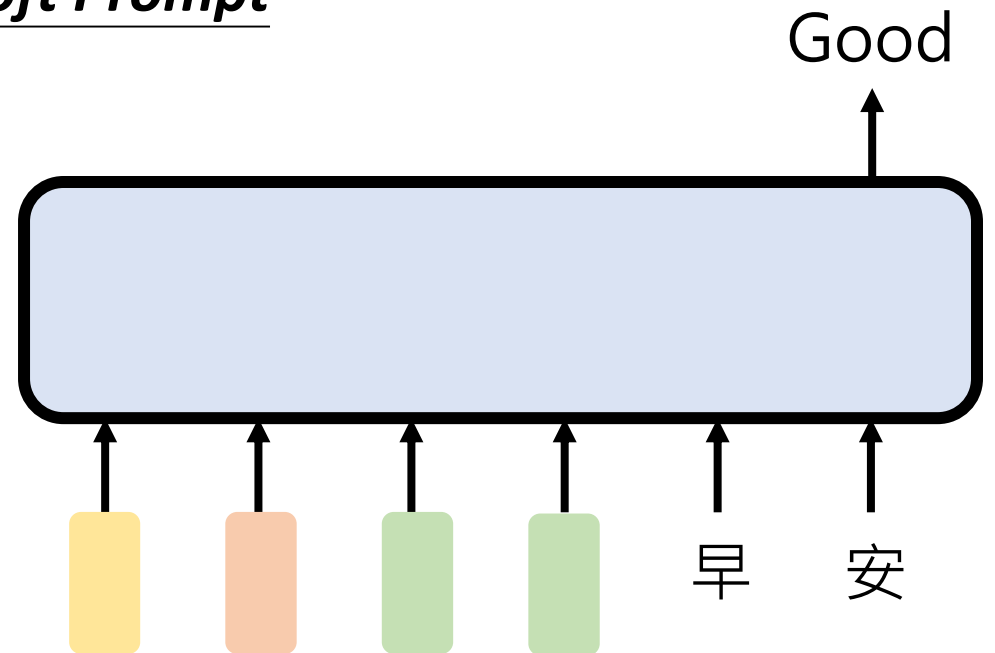
**A:** To answer the question "How many points did the Ravens have at halftime?", we need to know: "What were all the scores the Ravens had at halftime?"

# 用機器來找 Prompt

## Hard Prompt



## Soft Prompt



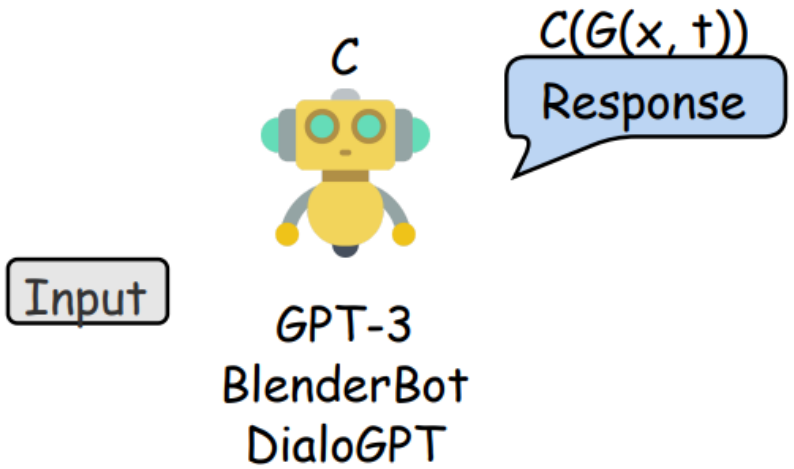
Trained by task-specific labeled data

Adapter at the input?



# 用機器來找 Prompt

- Using reinforcement learning



# 用機器來找 Prompt

- Using an LM to find prompt

今天天氣真好 分隔 正面 分隔

今天運氣真差 分隔 負面 分隔

這朵花真美 分隔 正面 分隔

我真的是累了 分隔 負面 分隔

## Forward Generation Template

I gave a friend an instruction and five inputs. The friend read the instruction and wrote an output for every one of the inputs. Here are the input-output pairs:

Input: [Q<sub>1</sub>]    Output: [A<sub>1</sub>]  
Input: [Q<sub>2</sub>]    Output: [A<sub>2</sub>]  
...

The instruction was **<COMPLETE>**

請決定這句話是正面還是負面

# 用機器來找 Prompt

## LLMs as Inference Models

Professor Smith was given the following instructions: **<INSERT>**

Here are the Professor's responses:

# Demonstration Start

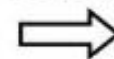
**Input:** prove    **Output:** disprove

**Input:** on      **Output:** off

...

# Demonstration End

Proposal



**write the antonym of the word.**

give the antonym of the word provided.

...

reverse the input.

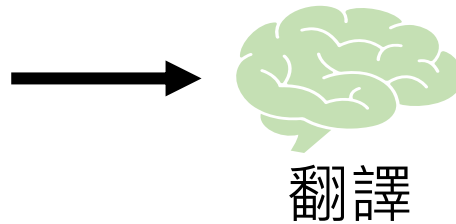
to reverse the order of the letters

No.	Category	Zero-shot CoT Trigger Prompt	Accuracy
1	APE	Let's work this out in a step by step way to be sure we have the right answer.	<b>82.0</b>
2	Human-Designed	Let's think step by step. (*1)	78.7
3		First, (*2)	77.3
4		Let's think about this logically.	74.5
5		Let's solve this problem by splitting it into steps. (*3)	72.2
6		Let's be realistic and think step by step.	70.8
7		Let's think like a detective step by step.	70.3
8		Let's think	57.5
9		Before we dive into the answer,	55.7
10		The answer is after the proof.	45.7
-		(Zero-shot)	17.7

# 兩種不同期待導致兩類不同的使用方式

- 期待一：成為專才

這堂課我們要講如何駕馭  
大型語言模型 .....

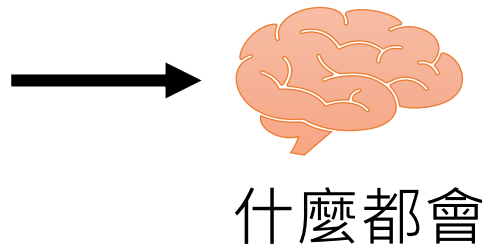


→ This course  
is about .....

- 期待二：成為通才

對以下文句做**翻譯**：

這堂課我們要講如何駕馭  
大型語言模型 .....



→ This course  
is about .....

# To learn more .....

- ACL 2022 Tutorial: Recent Advances in Pre-trained Language Models: Why Do They Work and How to Use Them
- Link: <https://d223302.github.io/AACL2022-Pretrain-Language-Model-Tutorial/>

