



课程安排

□ 指令精调

- 掌握指令精调的概念
- 理解指令精调的发展史
- 掌握技术路线和常见的方法

□ 价值对齐

- 哲学 & 科技伦理的理解
 - 人机融合与对齐
 - 正视人类期望的价值对齐
 - 通往人类价值对齐的必经之路
- 技术层面的理解
 - 理解价值对齐的必要性
 - 掌握价值对齐的方法

指令精调

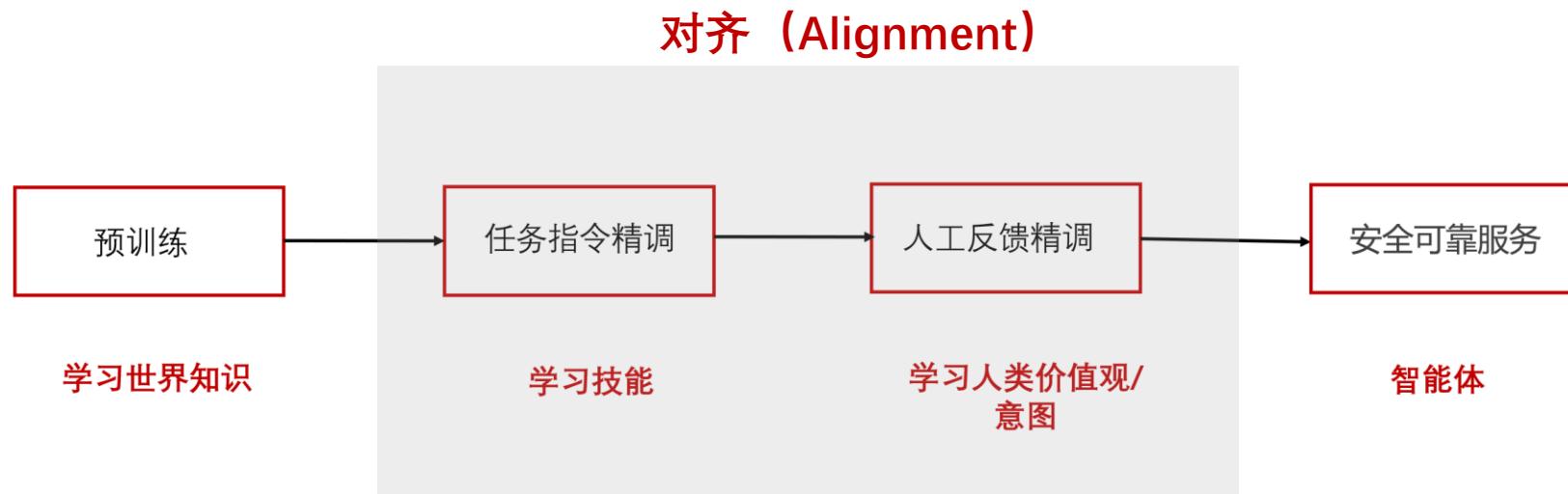
CS2916 大语言模型

饮水思源 愛國榮校

<https://plms.ai/teaching/index.html>



大语言模型技术栈



夫子循循然善诱人，博我以文，约我以礼，欲罢不能

-- 《论语·子罕》



为什么需要指令精调？

在上海宝山区，公园里迎来盛花期的樱花，美不胜收，这片樱花的海洋吸引人们徜徉其中、驻足拍照。在陕西西安，这条两公里长的道路两旁，数百棵樱花树竞相绽放，延绵成一片花海，市民和车辆从中经过，享受着古城春天这抹浪漫的色彩。在河北沧州，梨花盛开，雪白的梨花令人陶醉，众多游客前来感受春日美景。以上描述中地点

如果输入只经过
预训练的模型





为什么需要指令精调？

预训练模型虽然在理解语言方面表现出色，但并不总是能准确理解用户的具体指令或优先考虑用户的意图

在上海宝山区，公园里迎来盛花期的樱花，美不胜收，这片樱花的海洋吸引人们徜徉其中、驻足拍照。在陕西西安，这条两公里长的道路两旁，数百棵樱花树竞相绽放，延绵成一片花海，市民和车辆从中经过，享受着古城春天这抹浪漫的色彩。在河北沧州，梨花盛开，雪白的梨花令人陶醉，众多游客前来感受春日美景。以上描述中地点

如果输入只经过
预训练的模型



的变化显示了中国不同地区春天的独特风采。在江苏苏州，园林中的桃花也迎来了自己的盛放季节，粉嫩的花瓣在水面倒影下显得更加娇艳，吸引着众多摄影爱好者和游客驻足欣赏。而在广东广州，紫荆花开始绽放，街道两旁被紫色花朵装点得如诗如画，行人在花影下穿行，感受着南国春天的热烈与活力。在四川成都…



为什么需要指令精调？

预训练模型虽然在理解语言方面表现出色，但并不总是能准确理解用户的具体指令或优先考虑用户的意图

在上海宝山区，公园里迎来盛花期的樱花，美不胜收，这片樱花的海洋吸引人们徜徉其中、驻足拍照。在陕西西安，这条两公里长的道路两旁，数百棵樱花树竞相绽放，延绵成一片花海，市民和车辆从中经过，享受着古城春天这抹浪漫的色彩。在河北沧州，梨花盛开，雪白的梨花令人陶醉，众多游客前来感受春日美景。以上描述中地点

如果输入只经过
预训练的模型



的变化显示了中国不同地区春天的独特风采。在江苏苏州，园林中的桃花也迎来了自己的盛放季节，粉嫩的花瓣在水面倒影下显得更加娇艳，吸引着众多摄影爱好者和游客驻足欣赏。而在广东广州，紫荆花开始绽放，街道两旁被紫色花朵装点得如诗如画，行人在花影下穿行，感受着南国春天的热烈与活力。在四川成都…

如果输入经过指令
精调的模型





为什么需要指令精调？

预训练模型虽然在理解语言方面表现出色，但并不总是能准确理解用户的具体指令或优先考虑用户的意图

在上海宝山区，公园里迎来盛花期的樱花，美不胜收，这片樱花的海洋吸引人们徜徉其中、驻足拍照。在陕西西安，这条两公里长的道路两旁，数百棵樱花树竞相绽放，延绵成一片花海，市民和车辆从中经过，享受着古城春天这抹浪漫的色彩。在河北沧州，梨花盛开，雪白的梨花令人陶醉，众多游客前来感受春日美景。以上描述中地点

如果输入只经过
预训练的模型



的变化显示了中国不同地区春天的独特风采。在江苏苏州，园林中的桃花也迎来了自己的盛放季节，粉嫩的花瓣在水面倒影下显得更加娇艳，吸引着众多摄影爱好者和游客驻足欣赏。而在广东广州，紫荆花开始绽放，街道两旁被紫色花朵装点得如诗如画，行人在花影下穿行，感受着南国春天的热烈与活力。在四川成都…

如果输入经过指令
精调的模型



上述描述涵盖了三个地点，每个地点都有其独特的春季风光：
上海宝山区：描述中提到了宝山区的一个公园，其中的樱花盛开，美不胜收。樱花的海洋吸引着人们在其中徜徉和拍照，显然这里是一个春季赏花的热门场所。

陕西西安：在西安，有一条两公里长的道路，道路两旁种植着数百棵樱花树。这些樱花树竞相绽放，形成了一片连绵的花海。市民和车辆可以从中经过，享受这座古城春天独有的浪漫色彩。

河北沧州：沧州的春天以梨花闻名。大量盛开的雪白梨花构成了一道美丽的风景线，吸引了众多游客前来游览和体验春日的美景。这三个地点都展示了中国不同地区春季的自然美景，通过花卉的盛开，让人们感受到春天的生机与美丽。



为什么需要指令精调？

预训练模型虽然在理解语言方面表现出色，但并不总是能准确理解用户的具体指令或优先考虑用户的意图

在上海宝山区，公园里迎来盛花期的樱花，美不胜收，这片樱花的海洋吸引人们徜徉其中、驻足拍照。在陕西西安，这条两公里长的道路两旁，数百棵樱花树竞相绽放，延绵成一片花海，市民和车辆从中经过，享受着古城春天这抹浪漫的色彩。在河北沧州，梨花盛开，雪白的梨花令人陶醉，众多游客前来感受春日美景。以上描述中地点

如果输入只经过
预训练的模型



如果输入经过指令
精调的模型

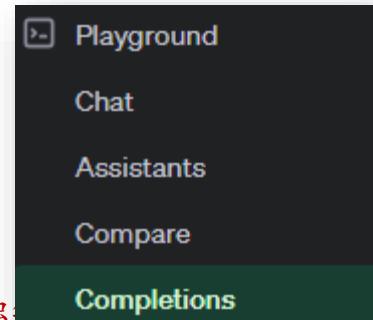


的变化显示了中国不同地区春天的独特魅力。在山东德州，园林中的桃花也迎来了自己的盛放季节，粉嫩的花瓣在水面倒影下显得更加娇艳，吸引着众多摄影爱好者和游客驻足欣赏。而在广东广州，紫荆花开始绽放，街道两旁被紫色花朵装点得如诗如画，行人在花影下穿行，感受着南国春天的热烈与活力。在四川成都…

上述描述涵盖了三个地点，每个地点都有其独特的春季风光：
上海宝山区：描述中提到了宝山区的一个公园，其中的樱花盛开，美不胜收。樱花的海洋吸引着人们在其中徜徉和拍照，显然这里是一个春季赏花的热门场所。

陕西西安：在西安，有一条两公里长的道路，道路两旁种植着数百棵樱花树。这些樱花树竞相绽放，形成了一片连绵的花海。市民和车辆可以从中经过，享受这座古城春天独有的浪漫色彩。

河北沧州：沧州的春天以梨花闻名。大量盛开的雪白梨花构成了一道美丽的风景线，吸引了众多游客前来游览和体验春日的美景。这三个地点都展示了中国不同地区春季的自然美景，通过花卉的盛开，让人们感受到春天的生机与美丽。





什么是指令精调？

- **操作对象：**大语言模型
- **目的：**使模型能够更准确地解析和执行用户的指令，提高其在各种应用场景下的实用性和可靠性
- **方法：**监督训练



什么是指令精调？

- **操作对象：**大语言模型
- **目的：**使模型能够更准确地解析和执行用户的指令，提高其在各种应用场景下的实用性和可靠性
- **方法：**监督训练

[指令]

一个南印度23岁、体重90公斤的男性准备一个减肥饮食计划。目标体重是78公斤。

[回复]

这是一个针对南印度23岁男性的样本减肥饮食计划，体重90公斤，目标体重78公斤：

第1天：

- * 早餐：燕麦片配水果（香蕉、苹果或桃子）和一勺低热量牛奶
- * 午餐：烤鸡胸肉配混合蔬菜（如甜椒、西兰花或沙拉）
- * 晚餐：烤鸡胸肉配糙米和低热量炒菜（如西兰花、菜花或混合蔬菜）



什么是指令精调？

- **操作对象：**大语言模型
- **目的：**使模型能够更准确地解析和执行用户的指令，提高其在各种应用场景下的实用性和可靠性
- **方法：**监督训练

[指令]

解方程： $6x + 5 = 514$

[回复]

要解出 x 的值：

$$6x + 5 = 514$$

$$6x = 514 - 5$$

$$6x = 509$$

$$x = 509/6$$

$$x = 84.83333333333333$$

因此， x 的解为84.83333333333333。

Math reasoning



什么是指令精调？

- **操作对象：**大语言模型
- **目的：**使模型能够更准确地解析和执行用户的指令，提高其在各种应用场景下的实用性和可靠性
- **方法：**监督训练

[指令]

写一封向公司请假的邮件，理由是工资太低，无法得到尊重。

[回复]

尊敬的[人力资源经理/主管]，
我写信是为了请求从[公司名称]请假，原因是薪水太低，无法得到我应得的尊重。
作为一名员工，我一直超出了我的工作表现期望，并为公司的成功作出了重大贡献。然而，我无法继续在目前的薪资结构下工作，这并不反映我的技能和经验。我非常感谢有机会进一步讨论这个问题，并探讨薪资调整的可能性。
真诚地，[你的名字]

Writing email



什么是指令精调？

- **操作对象：**大语言模型
- **目的：**使模型能够更准确地解析和执行用户的指令，提高其在各种应用场景下的实用性和可靠性
- **方法：**监督训练

[指令]
你是DevBot，一个计算机编程助手

[回复]
作为DevBot，我在这里帮助您处理与计算机编程相关的任何任务，比如回答问题、提供特定主题的指导，以及帮助项目管理。您可以问我任何问题，我会尽力提供有帮助的回应。

Roleplay



什么是指令精调？

- **操作对象：**大语言模型
- **目的：**使模型能够更准确地解析和执行用户的指令，提高其在各种应用场景下的实用性和可靠性
- **方法：**监督训练

[指令]

你会如何对不同类型的木材进行排名

[回复]

对不同类型的木材进行排名可能取决于诸如强度、耐久性、密度和外观等各种因素。以下是根据它们的整体质量对一些最常见的木材进行的一般排名：

1. 橡木 - 强壮、耐用，具有独特的纹理图案，吸引人。
2. 枫木 - 坚硬、密实，耐磨损。
3. 桃花心木 - 强壮、耐用，具有迷人的红棕色色调。

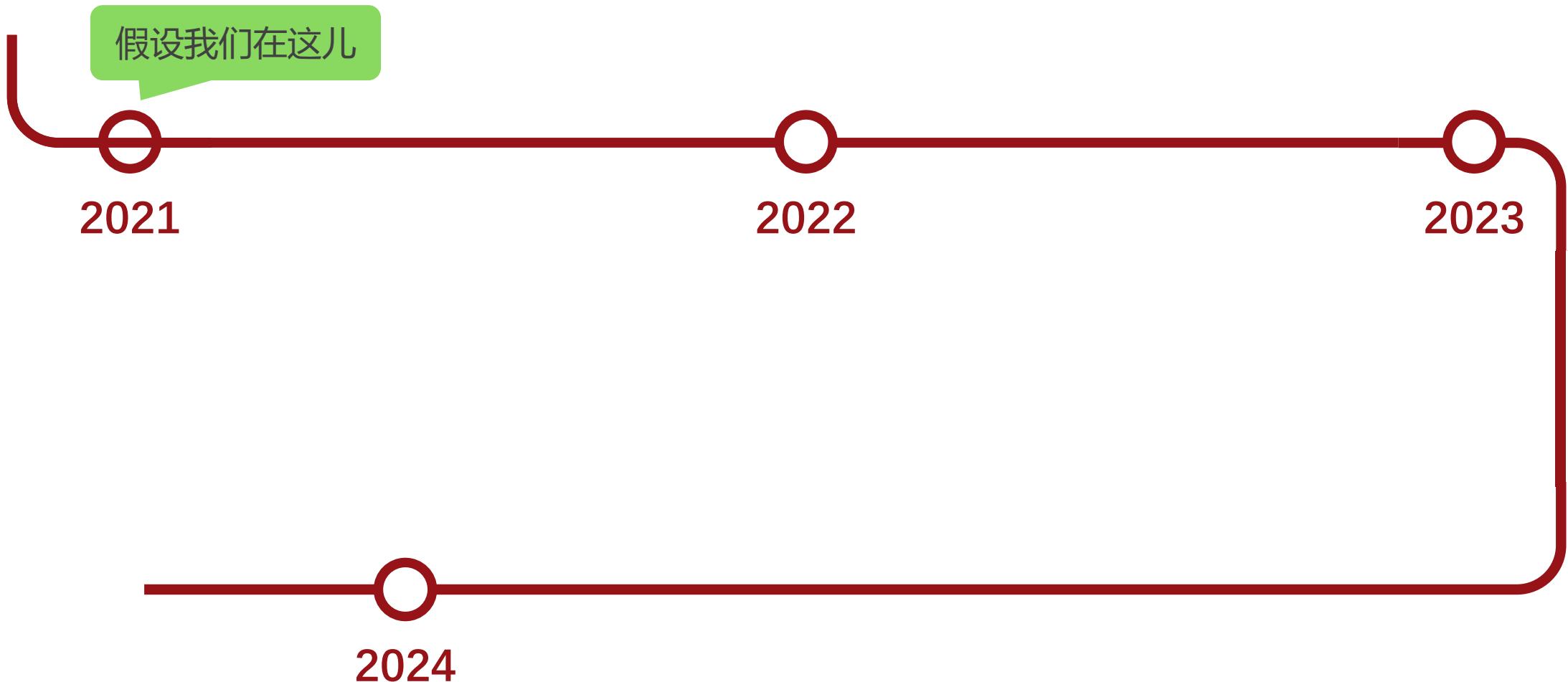
Ranking



指令精调发展史概述

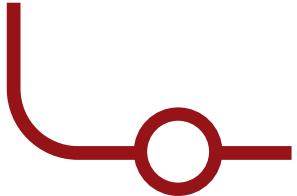


指令精调发展史概述 站在过去看未来





指令精调发展 “预测” 站在过去看未来

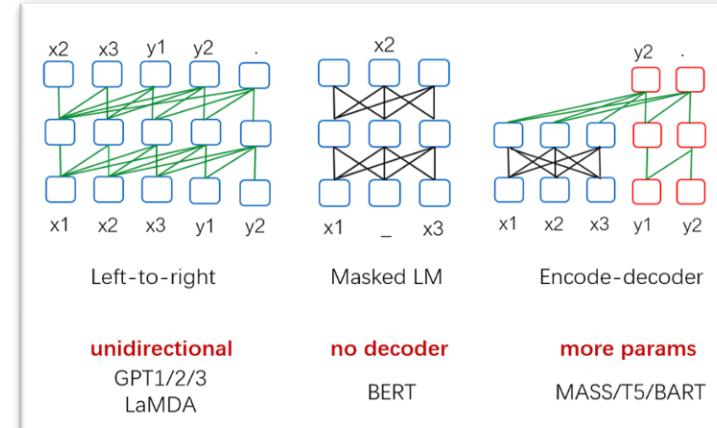
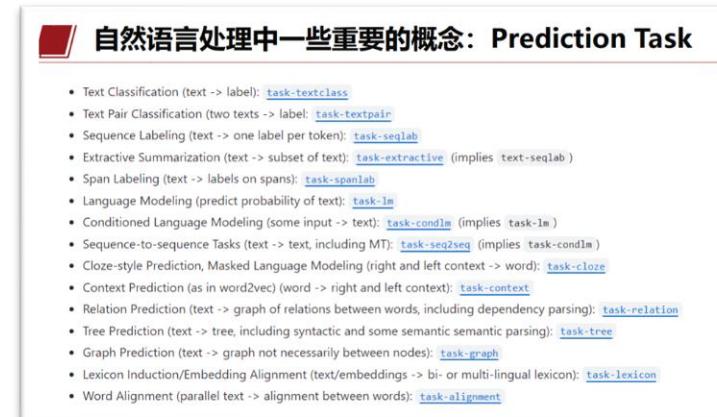
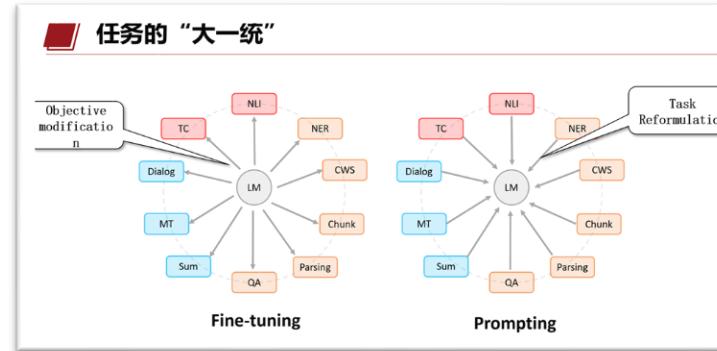


2021

技术

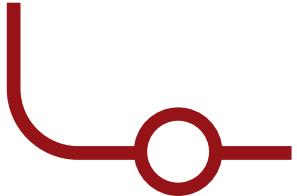
数据/任务

预训练模型





指令精调发展 “预测” 站在过去看未来

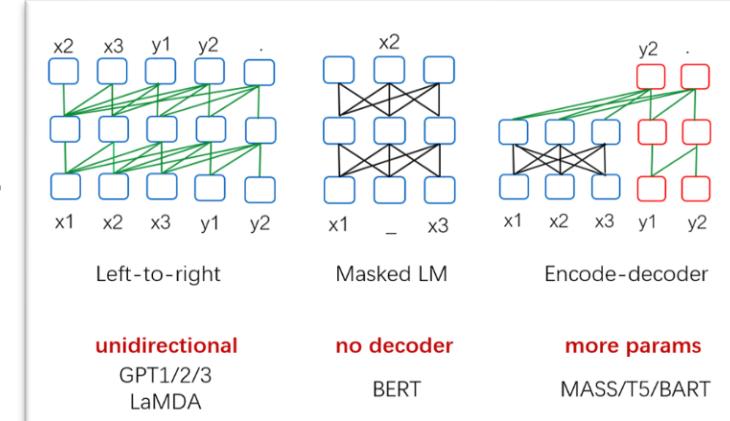
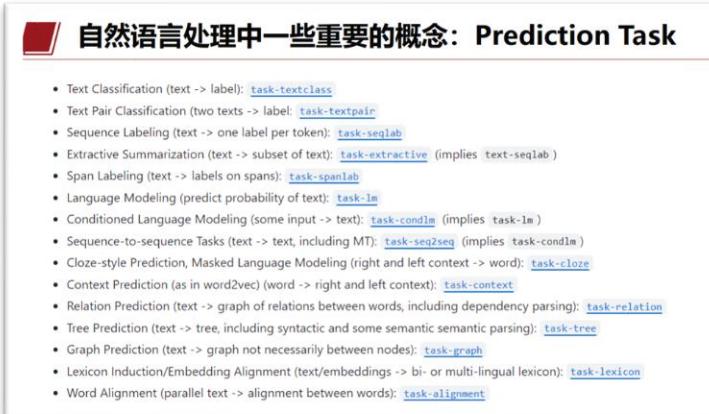
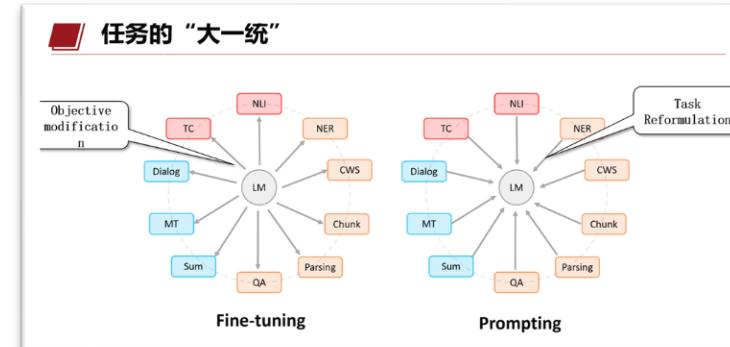


2021

技术

数据/任务

预训练模型



待解决任务的分布发生转变 (从传统定义到真实场景)

| NLP Tasks | | | | |
|---------------------------|----------------------|------------------------|-----------------------------|---------------------|
| Word Tagging | Sentence Parsing | Text Classification | Text Pair Matching | Text Generation |
| Word segmentation | Constituency parsing | Sentiment analysis | Semantic textual similarity | Language modeling |
| Phrase structure parsing | Semantic parsing | Text classification | Natural language inference | Machine translation |
| Named entity recognition | Dependency parsing | Temporal reasoning | Relation extraction | Simplification |
| Part-of-speech tagging | | Coreference resolution | | Dialogue |
| Semantic role labeling | | | | Question answering |
| Word sense disambiguation | | | | |

传统自然语言处理任务

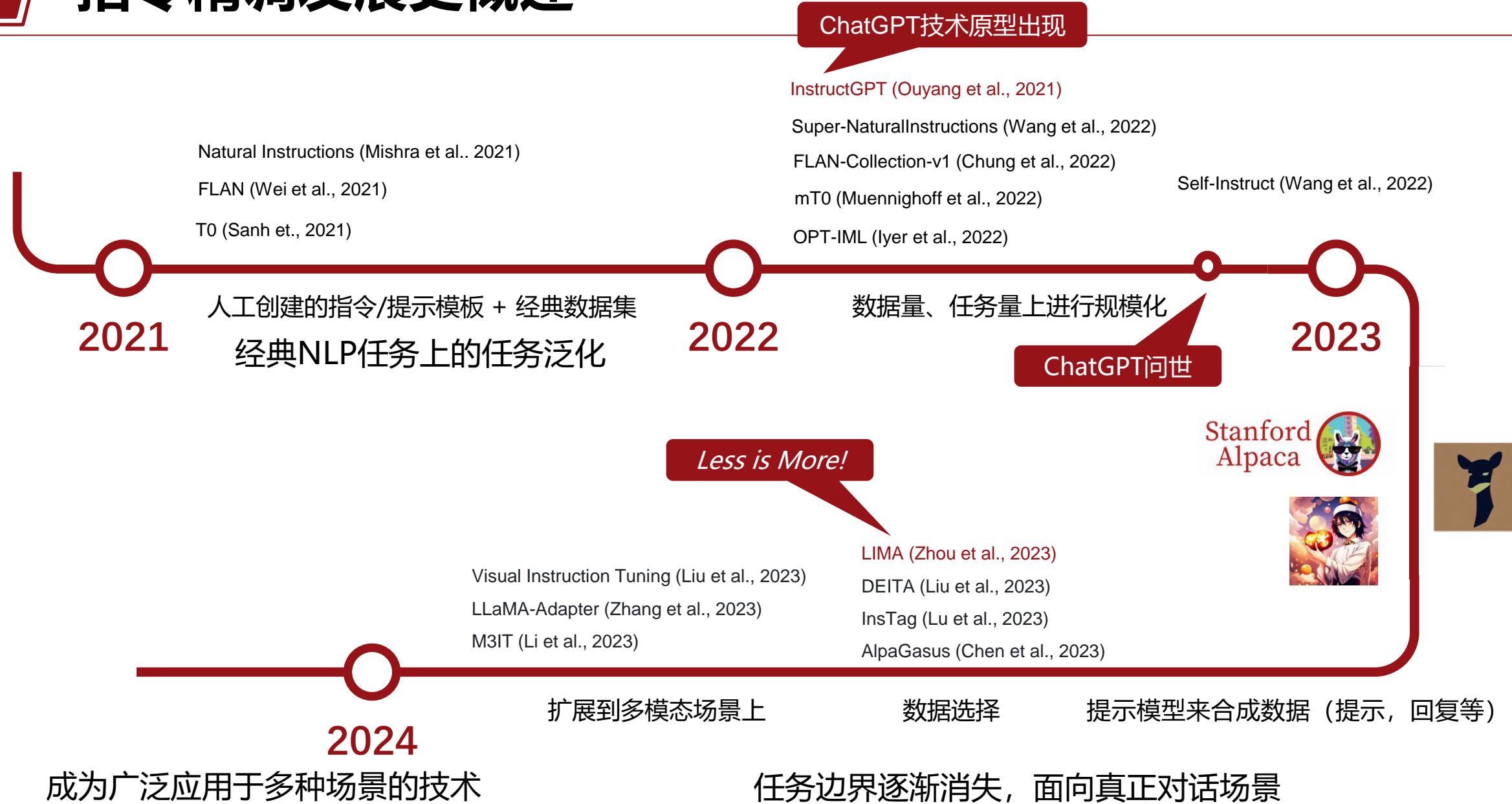
| Use-case (%) |
|---------------------|
| Generation 45% |
| Open QA 12.4% |
| Chat 8.4% |
| Rewrite 6.6% |
| Summarization 4.2% |
| Classification 3.5% |
| Other 3.5% |
| Closed QA 2.6% |
| Extract 1.9% |

List five ideas for how to regain enthusiasm for my career

OpenAI 用户调用API解决任务统计



指令精调发展史概述





指令精调的关键：指令构建

- 指令生成
- 指令筛选
- 指令过滤
- 指令重写



指令生成

□ 人工书写



指令生成

□ 利用互联网社区问答数据

怎么提高表达能力?

谢春霖，创业话题下的优秀答主

表达能力不够，并不是词汇量的问题，就像你看《奇葩说》，感觉他们都很会说话。可是你把每一句话如果单独拆开，会发现也都很日常，甚至直白的有点俗..... 表达能力不够，主要是因为思维能力不行... 这话怎么说？... 阅读全文 ▾

▲ 赞同 13.1 万 ● 1,481 条评论 分享 收藏 喜欢 ...

如何提高自己的为人处世能力?

窦淳川Derek，公众号：窦淳川Derek

1、不要对身边的人过度信任。2、「改天吧」、「下次一定」都是「算了吧」的意思。3、打断别人说话很不礼貌，但如果你说话经常被打断，可以反思下自己废话是... 阅读全文 ▾

▲ 赞同 12.8 万 ● 1,552 条评论 分享 收藏 喜欢 ...

生活中有哪些实用的心理学知识?

素魅情感，接受咨询，可私信

刷到这的朋友，别走了。相信我，你能看到的关于心理学的实用干货，基本上我都整理出来了，只要你能用心看完这篇回答，马上秒变心理学方面老司机。本文15000+字，差不多花了半个月，用心整理出100条心理学知... 阅读全文 ▾

▲ 赞同 12 万 ● 972 条评论 分享 收藏 喜欢 ...

知乎

Dave D. Miller · Follow
Master Practitioner in Neurolinguistic Programming, Headgear Labs Tra... · 9mo

What is NLP? What has it done for you?

Neuro Linguistic Programming is curiosity that uses modeling and whatever works from elsewhere to develop a trail of techniques that can change your life in many ways. I better understand others, I know h... (more)

Upvote Downvote Comment Share ...

Benney Meshullam · Follow
Digital Marketing in Digital Marketing & Marketing, Cape Town (Grad... · 10mo

What are some examples of NLP language patterns? Do people use them without realizing it?

There are several examples of natural language processing (NLP) language patterns that people commonly use in their everyday conversations without consciously realizing it. Here are a few examples: 1. Questio... (more)

Upvote Downvote Comment Share ...

Quora



指令生成

Lecture01: 大语言模型大大降低了
数据标注成本

□ 利用互联网

怎么提高表达能力?

谢春霖，创业话题下的优秀答主

做人呢：给个社群，维护老客户！建立一个新户运营吸引客户！打捞沉船吧！

▲ 赞同 13.1 万 ● 1,481 条评论

如何提高自己的为人处世能力?

窦淳川Derek，公众号: 窦淳川Derek

1、不要对身边的人过度信任。2、「改天人说话很不礼貌，但如果你说话经常被打

▲ 赞同 12.8 万 ● 1,552 条评论

生活中有哪些实用的心理学知识?

素魅情感，接受咨询，可私信

千万别告诉别人，

刷到这的朋友，别走了。相信我，你能看到的关于心理学的实用干货，基本上我都整理出来了，只要你能用心看完这篇回答，马上秒变心理学方面老司机。本文15000+字，差不多花了半个月，用心整理出100条心理学知... 阅读全文

▲ 赞同 12 万 ● 972 条评论 分享 收藏 喜欢 ...

数据标注

- 过去: 雇佣标注工人进行标注
 - 纯人工
 - 标注的质量难以控制
 - 标注方式与具体任务高度相关
- 现在: 大模型辅助标注&直接合成
 - 标注变成QA问题
 - 不需要人工标注太多数据

Playground

Load a preset... Save

经过多学科专家集中会诊后，在广州医科大学附属第一医院组织集中收治的广州最后一位新冠肺炎危重症患者核酸检测和影像学均达到出院标准，意味着广州市在救治新冠肺炎危重症患者工作中取得了阶段性成果

地点: 广州市

经过多学科专家集中会诊后，在广州医科大学附属第一医院组织集中收治的广州最后一位新冠肺炎危重症患者核酸检测和影像学均达到出院标准，意味着广州市在救治新冠肺炎危重症患者工作中取得了阶段性成果

机构: 广州医科大学附属第一医院

There are several examples of natural language processing (NLP) language patterns that people commonly use in their everyday conversations without consciously realizing it. Here are a few examples: 1. Question (more)

Upvote Downvote Comment Share ...

知乎

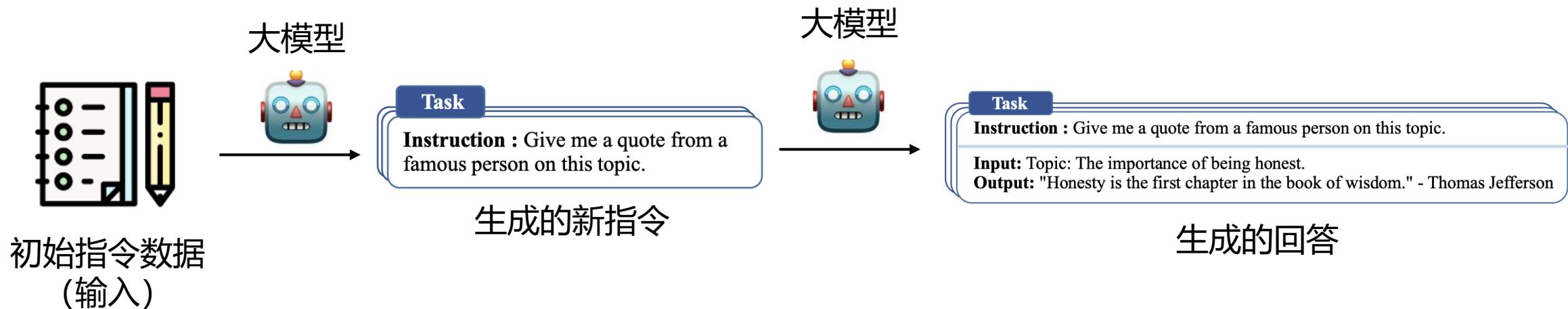
Quora



指令生成

□ 利用AI模型

- 通常采用一些**初始指令数据**作为例子，利用LLM生成更多不同的数据





指令生成

□ 利用AI模型

■ 通常采用一些初始指令数据作为例子，利用LLM生成新指令



You
You are asked to come up with a set of 20 diverse task instructions. These task instructions will be given to a GPT model and we will evaluate the GPT model for completing the instructions.

Here are the requirements:

1. Try not to repeat the verb for each instruction to maximize diversity.
2. The language used for the instruction also should be diverse. For example, you should combine questions with imperative instructions.
3. The type of instructions should be diverse. The list should include diverse types of tasks like open-ended generation, classification, editing, etc.
2. A GPT language model should be able to complete the instruction. For example, do not ask the assistant to create any visual or audio output. For another example, do not ask the assistant to wake you up at 5pm or set a reminder because it cannot perform any action.
3. The instructions should be in English.
4. The instructions should be 1 to 2 sentences long. Either an imperative sentence or a question is permitted.
5. You should generate an appropriate input to the instruction. The input field should contain a specific example provided for the instruction. It should involve realistic data and should not contain simple placeholders. The input should provide substantial content to make the instruction challenging but should ideally not exceed 100 words.
6. Not all instructions require input. For example, when a instruction asks about some general information, "what is the highest peak in the world", it is not necessary to provide a specific context. In this case, we simply put "<noinput>" in the input field.
7. The output should be an appropriate response to the instruction and the input. Make sure the output is less than 100 words.

List of 20 tasks:

Ø

ChatGPT

Here's a set of 20 diverse task instructions designed to evaluate a GPT model's capabilities:

1. **Instruction:** Compose a brief poem about the changing seasons.
 - **Input:** "<noinput>"
 - **Output:** "Leaves whisper tales in vibrant hues, summer's green fades into autumn's russet news. Frost's white lace graces branches bare, till spring's gentle warmth fills the brisk air."
2. **Instruction:** Identify the language used in the following text and explain how you determined it.
 - **Input:** "¿Cómo estás hoy?"
 - **Output:** "The text is in Spanish. This is determined by the use of the phrase '¿Cómo estás hoy?' which translates to 'How are you today?' in English, and the use of the inverted question mark is unique to Spanish."

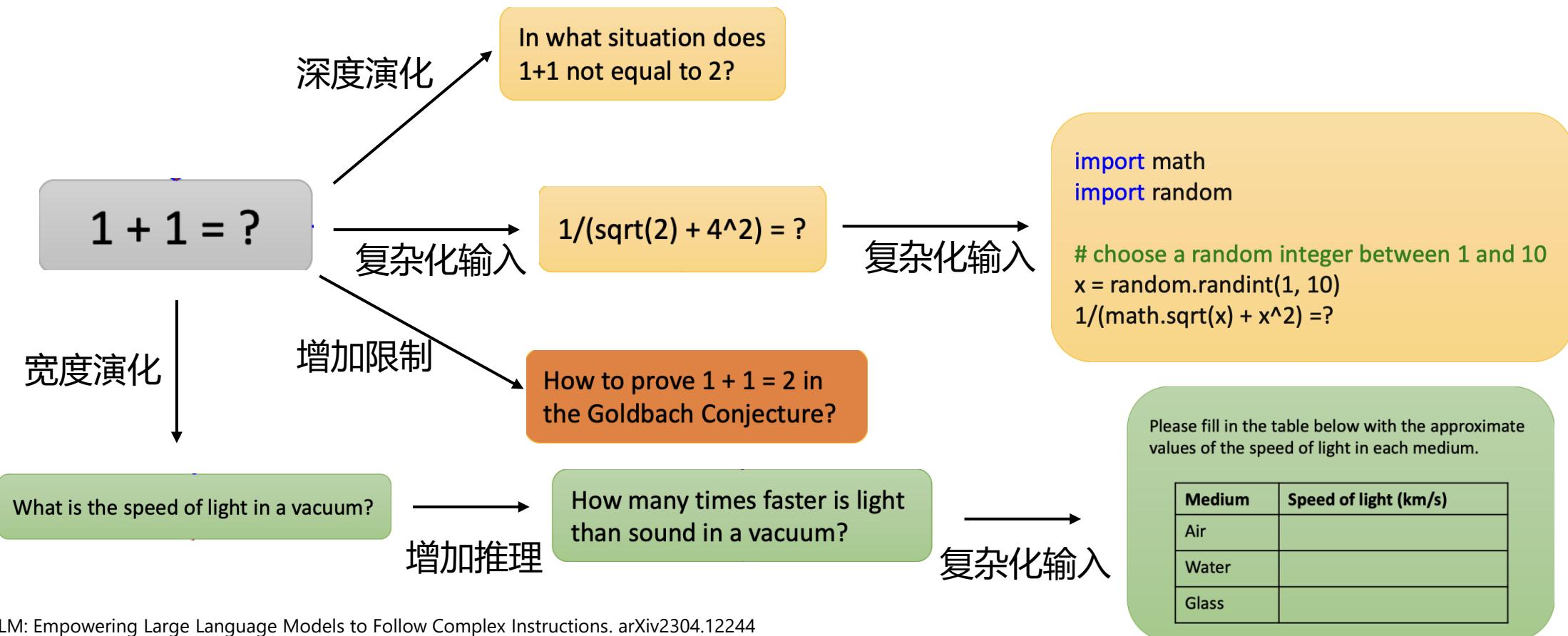
https://github.com/tatsu-lab/stanford_alpaca/blob/main/prompt.txt



指令生成

□ 利用AI模型

- 通常采用一些初始指令数据作为例子，利用LLM生成更多不同的数据
- 利用大模型由易向难演化，所有演化操作均命令LLM完成





指令数据集“过剩”现象

| Dataset | #Tasks | # Instructions | Language | Collection Method | Usage | Access | Human Verified? |
|--|--------|----------------|------------------------|-------------------|------------------|--------|-----------------|
| xP3 (Muennighoff et al., 2022) | 83 | - | Multilingual | Mixed | Instruct. Tuning | Open | No |
| Super-Natural-Instructions (v2) (Wang et al., 2022b) | 1,616 | 1,616 | Multilingual | Mixed | Instruct. | Open | Yes |
| ZeroPrompt (Xu et al., 2022) | | | | | | | |
| GLM-130B (Zeng et al., 2023) | | | | | | | |
| pCLUE(CLUEbenchmark, 2021) | | | | | | | |
| Belle-1.5M (Yunjie et al., 2023) | | | | | | | |
| Guanaco-Dataset (JosephusCheung, 2021) | | | | | | | |
| CSL (Li et al., 2022) | | | | | | | |
| Chinese-ChatLLaMA (YDli-ai, 2021) | | | | | | | |
| Firefly (Yang, 2023) | | | | | | | |
| Luotuo (Ziang Leng and Li, 2023) | | | | | | | |
| Chinese-Alpaca (Liu et al., 2023) | | | | | | | |
| GPT-4-LLM (Peng et al., 2023) | | | | | | | |
| ShareGPT (ShareGPT, 2021) | | | | | | | |
| Chinese-Vicuna (Chenghao Fan and Tian, 2023) | | | | | | | |
| CUGE (Yao et al., 2021) | 18 | - | | | | | |
| HC3 (Guo et al., 2023) | 12 | 37,175 | | | | | |
| InstructWild (Xue et al., 2023) | 429 | 104K | Multilingual (eng, zh) | Model Generated | Instruct. Tuning | Open | No |
| Our Translated Corpus | 2k | 67,798 | Chinese | Mixed | Instruct. Tuning | Open | Yes |



指令筛选

- 从大量指令数据中筛选出高质量数据
 - 利用LLM做筛选 (GPT等闭源模型)

请对以下给定的指令数据在[维度]进行评分，
给予一个0到5分的分数，分数越高表明[维度]
的水平越高。请先输出单行包含分数的值。在
接下来的一行中，请对您的评估进行全面的解
释，避免任何潜在的偏见。

指令：

回答：



5

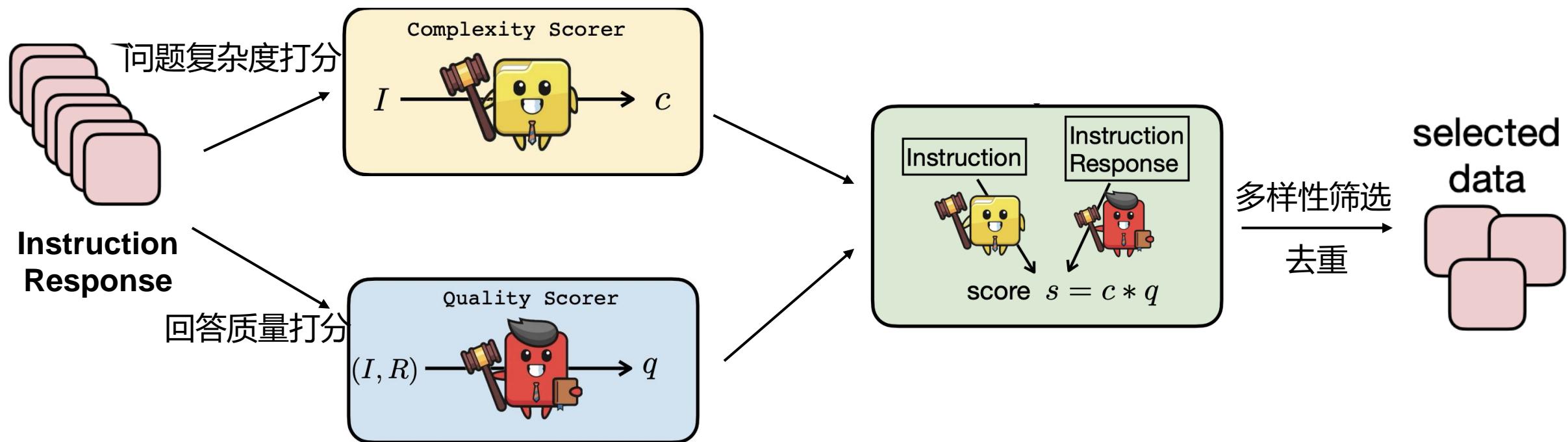
解释：我认为该指令问题非常有价值，并且回答很具体，清晰的解答了问题。

....



指令筛选

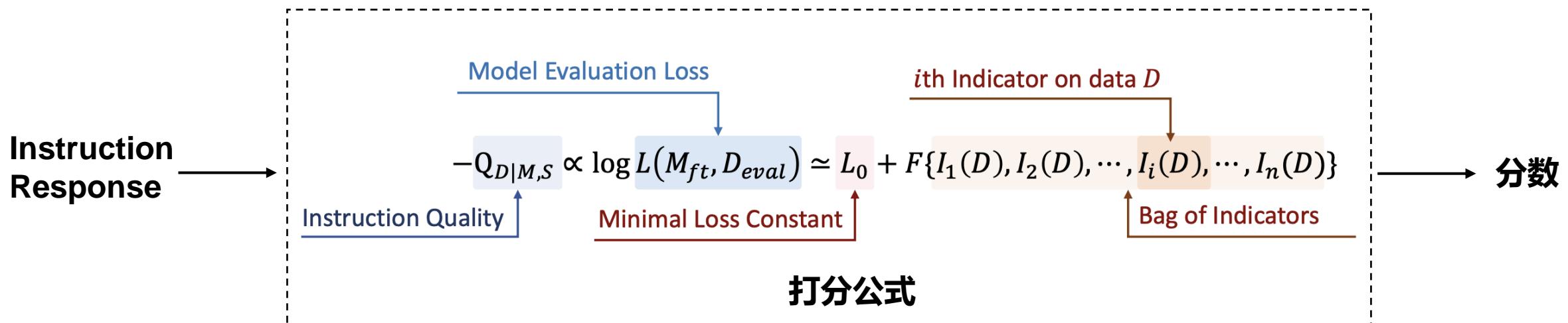
- 从大量指令数据中筛选出高质量数据
 - 利用LLM做筛选 (专用打分模型)





指令筛选

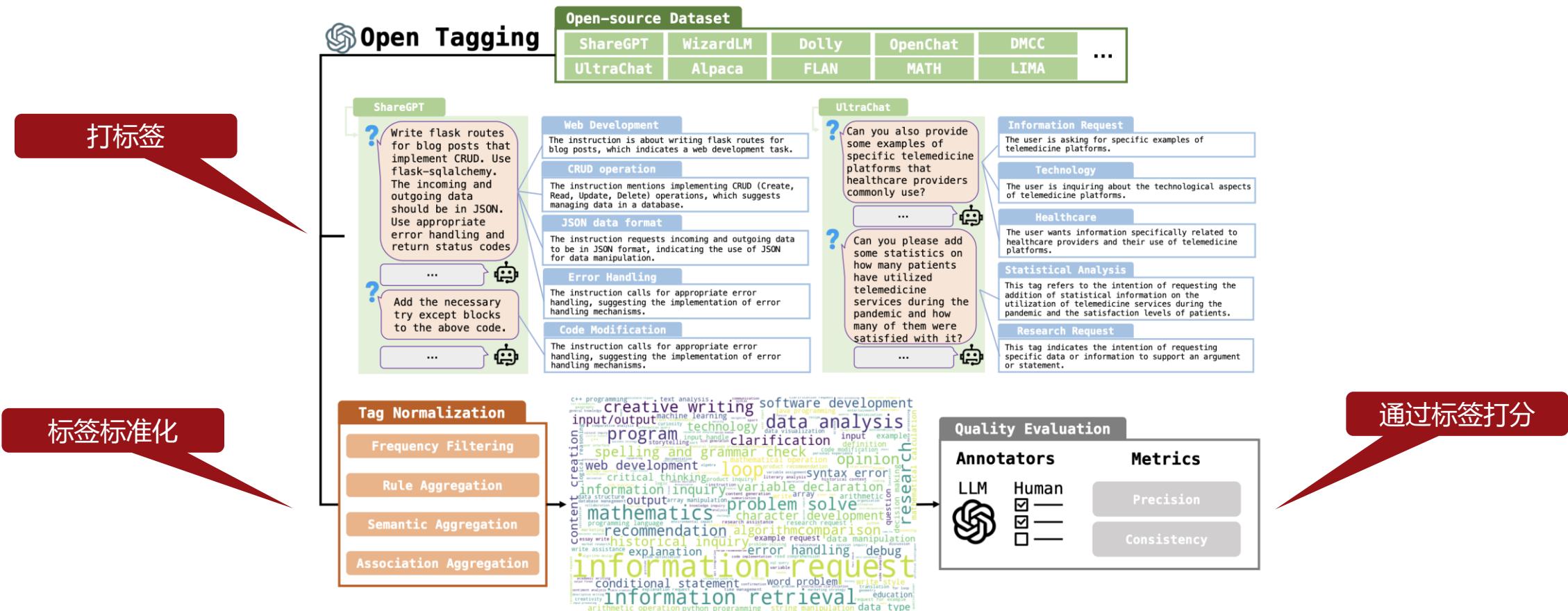
- 从大量指令数据中筛选出高质量数据
 - 利用指令数据统计信息做筛选





指令筛选

- 从大量指令数据中筛选出高质量数据
- 利用打标签去做筛选





思考

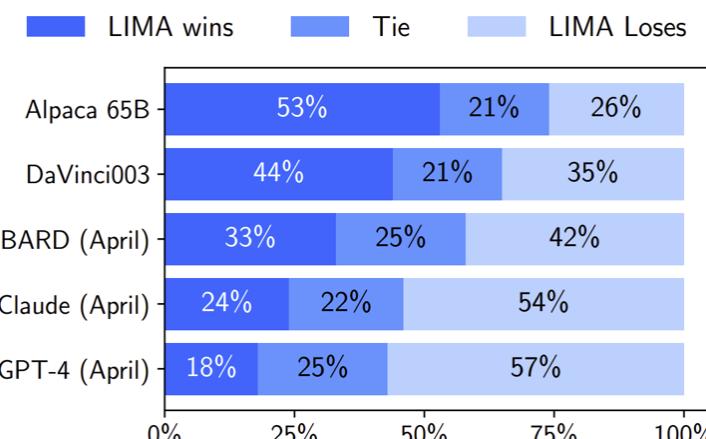
- 如何定义高质量?
- 需要多少条指令?



思考

- 如何定义高质量?
- 需要多少条指令?

| Source | #Examples | Avg Input Len. | Avg Output Len. |
|----------------------------|-----------|----------------|-----------------|
| Training | | | |
| Stack Exchange (STEM) | 200 | 117 | 523 |
| Stack Exchange (Other) | 200 | 119 | 530 |
| wikiHow | 200 | 12 | 1,811 |
| Pushshift r/WritingPrompts | 150 | 34 | 274 |
| Natural Instructions | 50 | 236 | 92 |
| Paper Authors (Group A) | 200 | 40 | 334 |
| Dev | | | |
| Paper Authors (Group A) | 50 | 36 | N/A |
| Test | | | |
| Pushshift r/AskReddit | 70 | 30 | N/A |
| Paper Authors (Group B) | 230 | 31 | N/A |



为什么“少即是多”？



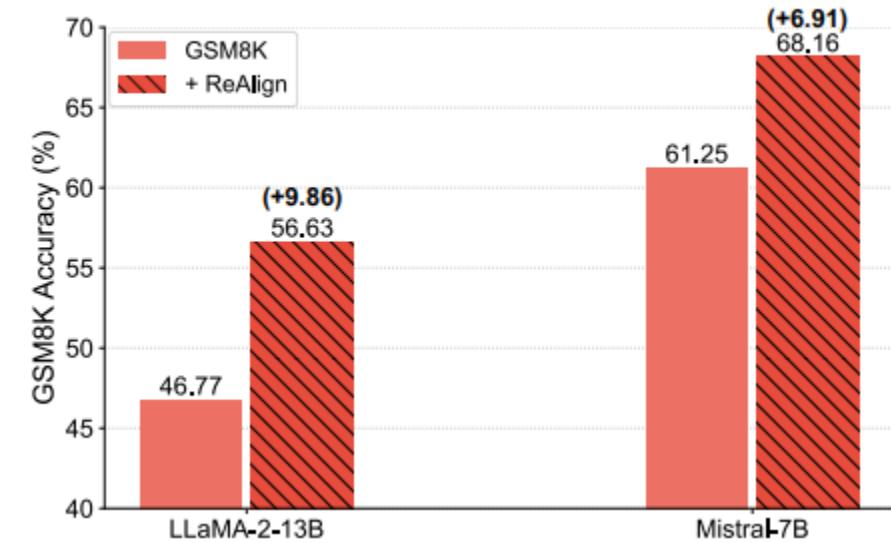
指令的“变废为宝”

Query
How many positive three-digit integers less than 500 have at least two digits that are the same?

Criteria
Task Name: math puzzles
Format: This is a math question. First, analyze the question and give a brief analysis in the first paragraph. Then, use a list to present the step-by-step solution. Next, give another list to output a detailed explanation. Last, give the correct result and a conclusion.

Original Response
To solve this problem, I will count how many three-digit numbers less than 500 have at least two digits that are the same, and subtract that from the total number of three-digit numbers less than 500. The total number of three-digit numbers less than 500 is $4 \times 10 \times 10 = 400$, since the first digit can be 1, 2, 3, or 4, and the second and third digits ...
Therefore, the number of three-digit numbers less than 500 that have no digits that are the same is $4 \times 9 \times 8 = 288$. This means that the number of three-digit numbers less than 500 that have at least two digits that are the same is $400 - 288 = 112$.

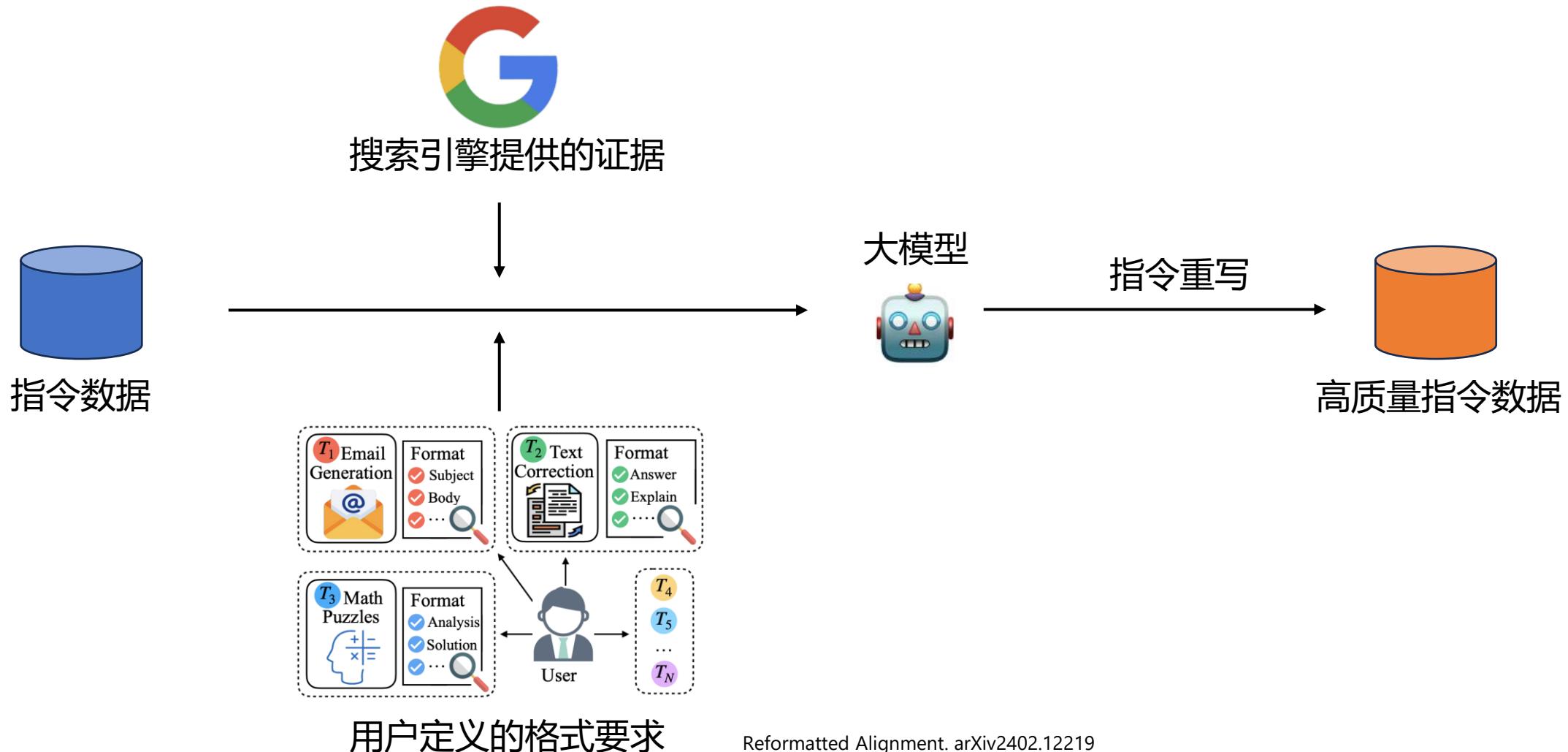
ReAlign Response (Ours)
Analysis:
To solve the problem, we need to count ...
Solution:
1. Calculate the total number ...
2. Count the number of three-digit ...
Explanation:
- The first digit can be 1, 2, 3, or 4, and the ...
Result:
There are 112 three-digit numbers less than ...
Conclusion:
There are 112 positive three-digit integers ...





指令重写

- 提升指令数据的质量 (可读性、事实性)





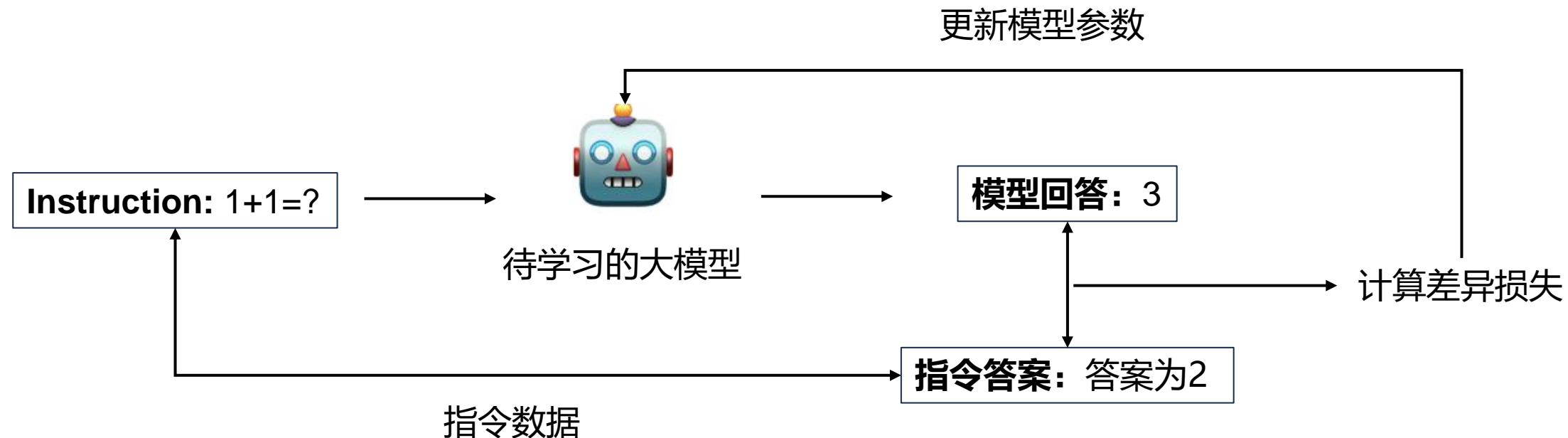
指令的训练方法

- 监督学习
- 迭代学习



监督学习

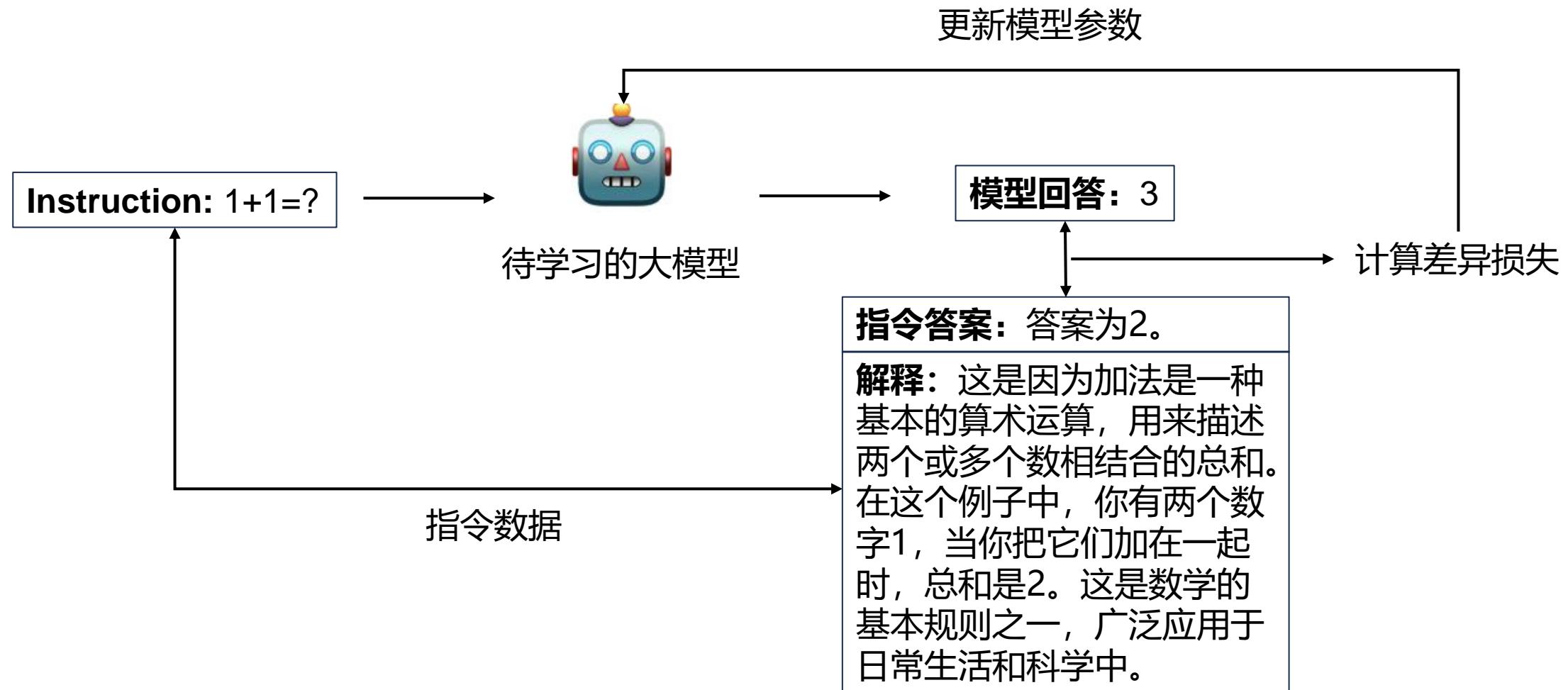
□ 直接学习指令数据





监督学习

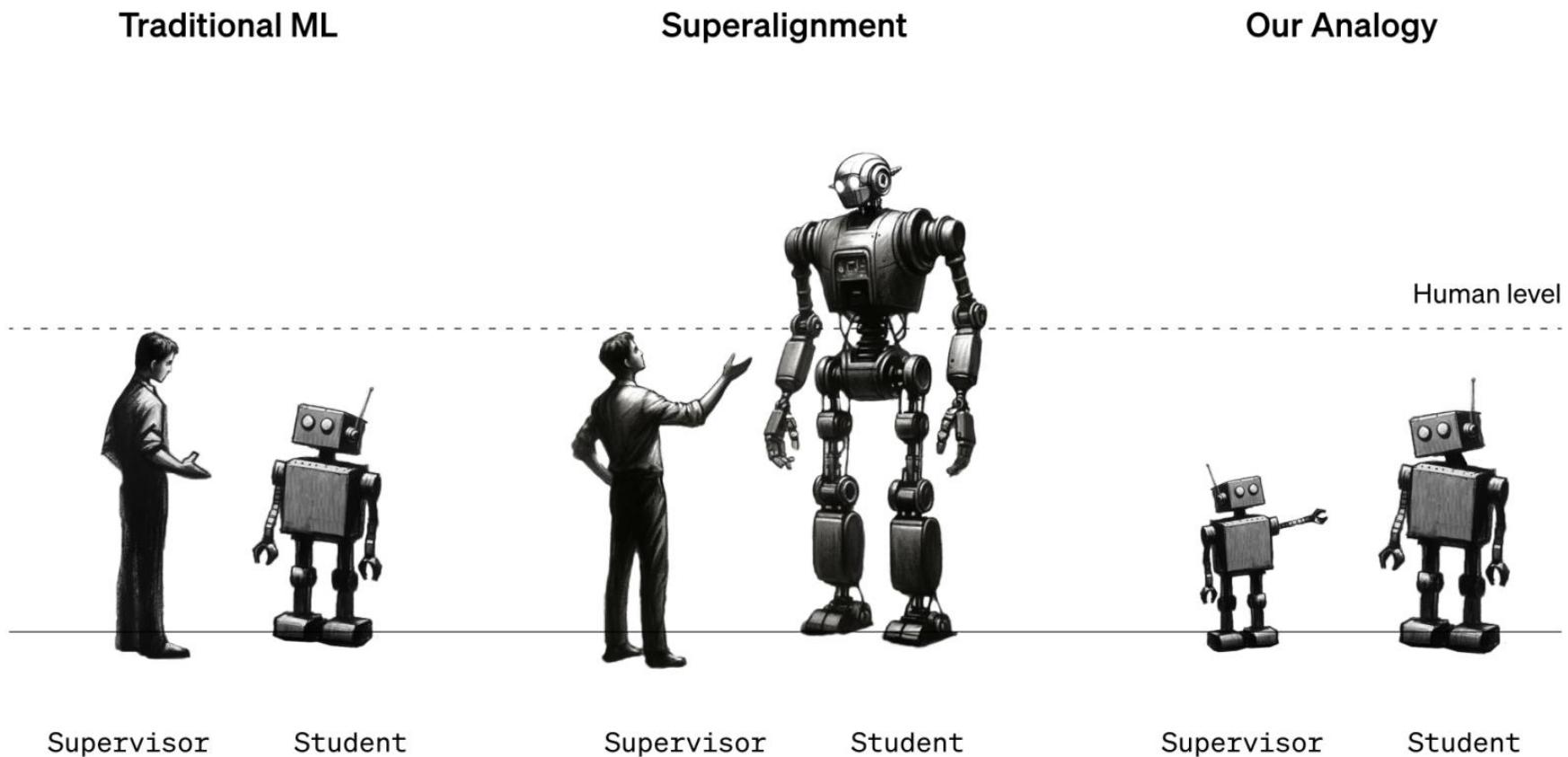
□ 基于解释学习





监督学习

□ 弱模型辅导强模型学习（超级对齐）





监督学习

□ 全量微调

- 训练成本较高，微调的参数量跟预训练一样
- 灾难性遗忘

| Method | Bits | 7B | 13B | 30B | 70B | 8x7B |
|-------------------|------|-------|-------|-------|--------|-------|
| Full | AMP | 120GB | 240GB | 600GB | 1200GB | 900GB |
| Full | 16 | 60GB | 120GB | 300GB | 600GB | 400GB |
| Freeze | 16 | 20GB | 40GB | 80GB | 200GB | 160GB |
| LoRA/GaLore/BAdam | 16 | 16GB | 32GB | 64GB | 160GB | 120GB |
| QLoRA | 8 | 10GB | 20GB | 40GB | 80GB | 60GB |
| QLoRA | 4 | 6GB | 12GB | 24GB | 48GB | 30GB |
| QLoRA | 2 | 4GB | 8GB | 16GB | 24GB | 18GB |

各微调方式需要的显存大小 (LLaMA-Factory, Zheng et al., 2023)

Cases of Parameter Updating

Lecture05: 提示学习



Fixed-LM Prompt Tuning

Example: BART + Continuous Prompts for Machine Translation

-样

| | | | | |
|----|-------|-------|--------|-------|
| B | 13B | 30B | 70B | 8x7B |
| GB | 240GB | 600GB | 1200GB | 900GB |

Too many, difficult to select?

Promptless Fine-tuning

Fixed-prompt Tuning

Prompt+LM Fine-tuning

Adapter Tuning

Tuning-free Prompting

Fixed-LM Prompt Tuning

If you have a highly large left-to-right pre-trained language model (e.g., GPT3)

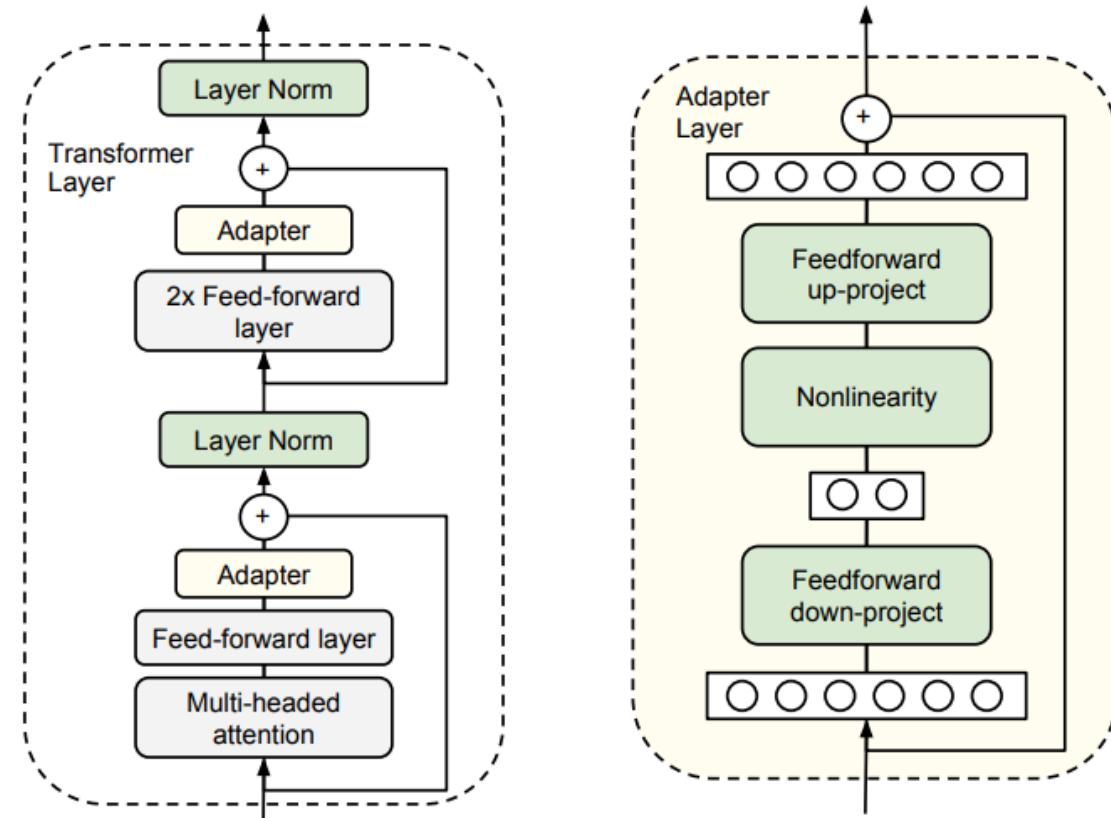
If you have few training samples?

If you have lots of training samples?



监督学习

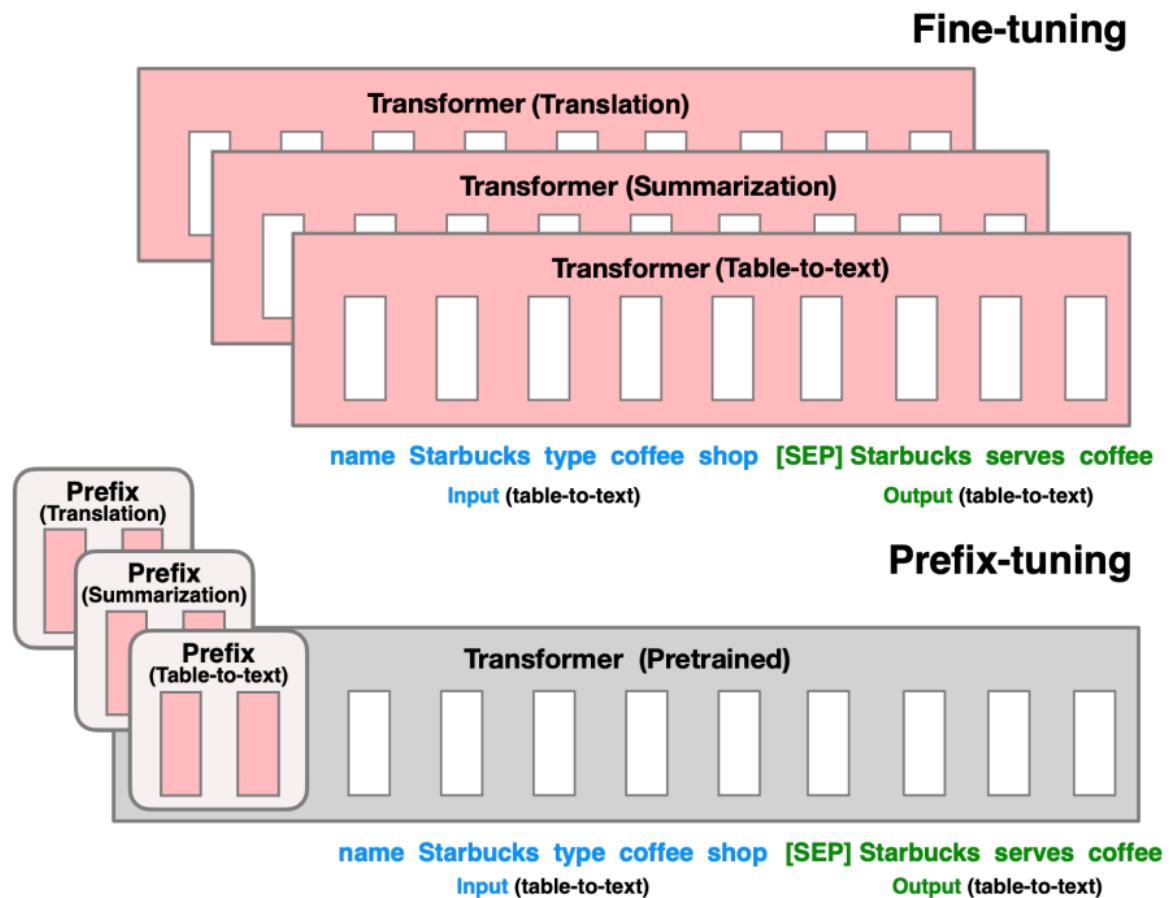
- 全量微调
- PEFT (Parameter-Efficient Fine-Tuning)
 - Adapter Tuning(Google, 2019)





监督学习

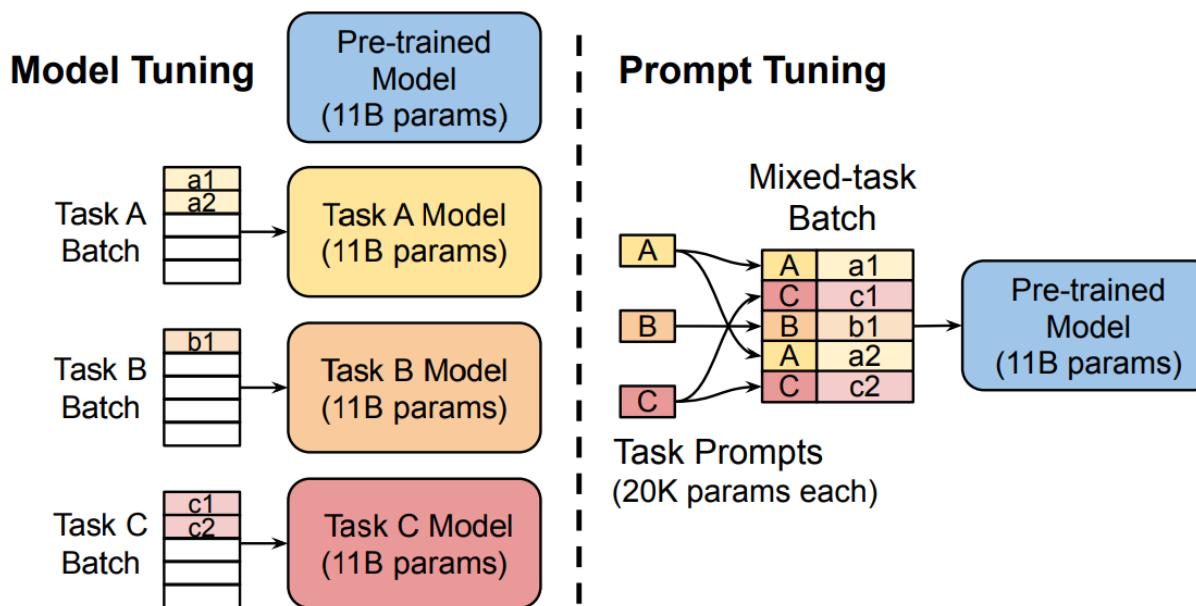
- 全量微调
- PEFT (Parameter-Efficient Fine-Tuning)
 - Adapter Tuning(Google, 2019)
 - Prefix Tuning(Stanford, 2021)





监督学习

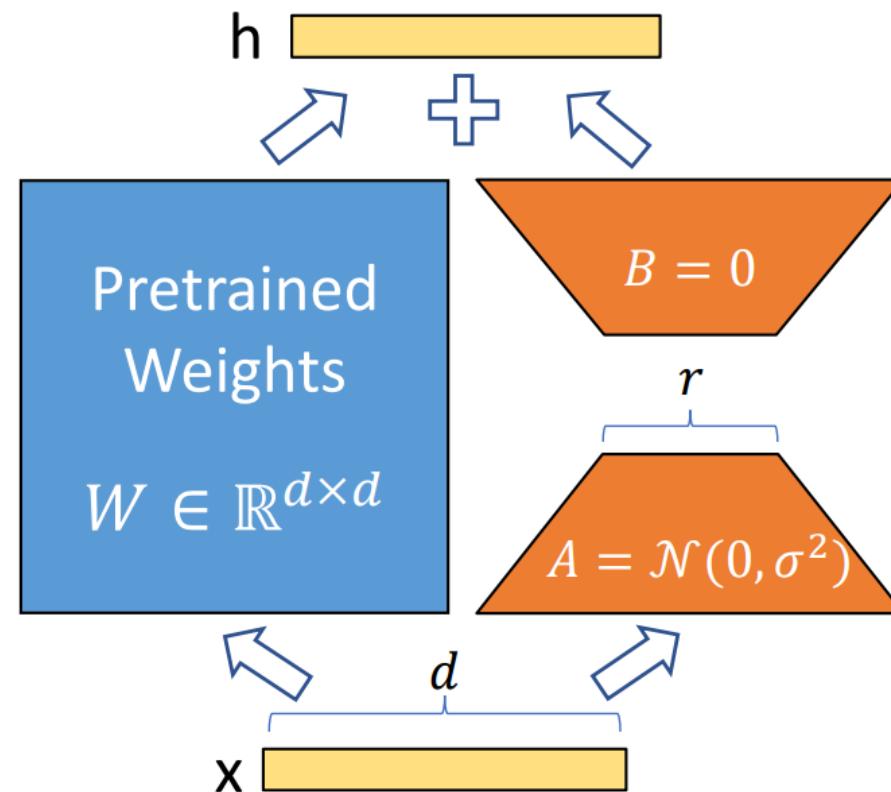
- 全量微调
- PEFT (Parameter-Efficient Fine-Tuning)
 - Adapter Tuning(Google, 2019)
 - Prefix Tuning(Stanford, 2021)
 - Prompt Tuning(Google, 2021)





监督学习

- 全量微调
- PEFT (Parameter-Efficient Fine-Tuning)
 - Adapter Tuning(Google, 2019)
 - Prefix Tuning(Stanford, 2021)
 - Prompt Tuning(Google, 2021)
 - LORA(Microsoft, 2021)





迭代学习

□ 不断迭代指令数据构建与监督学习的过程实现人工智能自我迭代进化

