

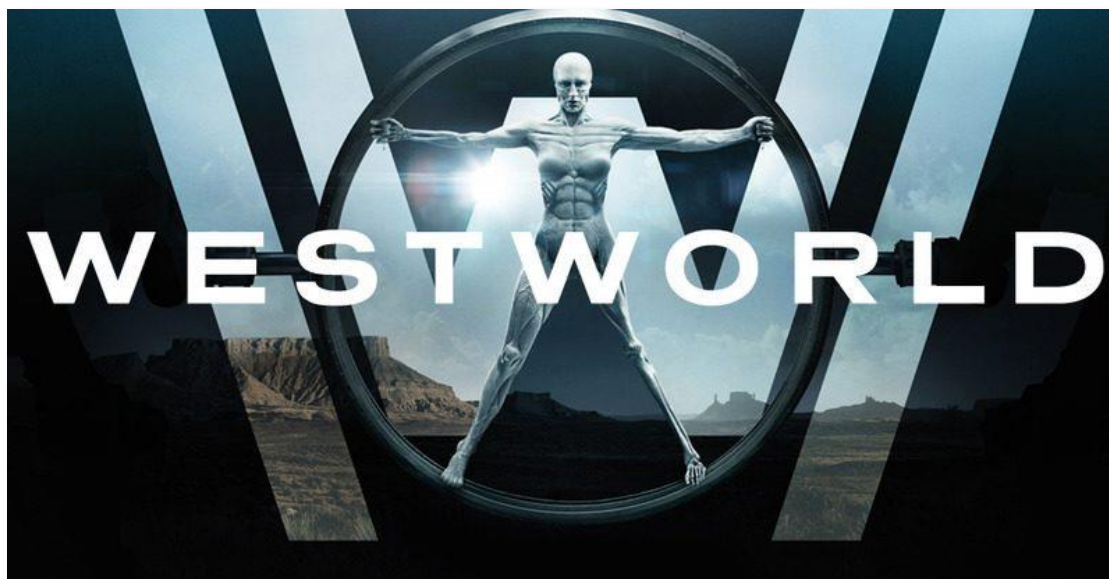
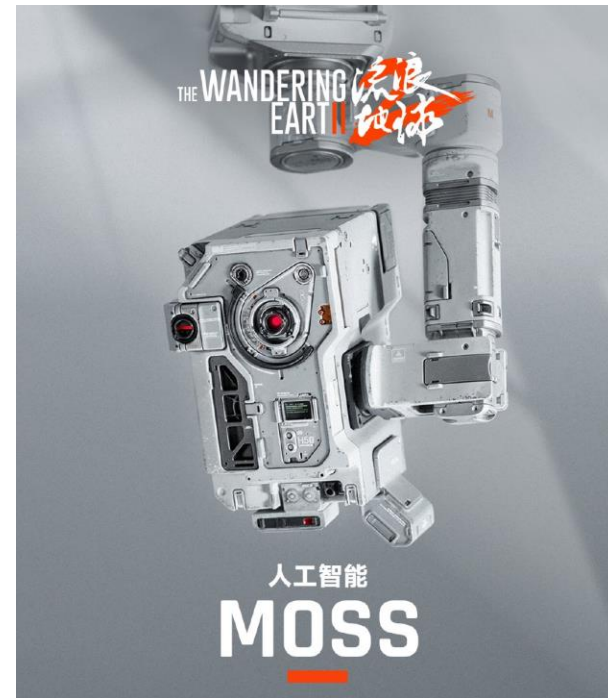


大语言模型理论与实践

张奇

复旦大学

ChatGPT仿佛电影中的AI就在眼前



ChatGPT给公众展现出来的是无所不能

GM

选择题，请运用你的中国哲学知识选择正确的答案，并说出你推理的过程。汉晋时期有多种文本记载，帝尧之时，“天下太和，百姓无事”。有老者“击壤”而戏，围观者称颂帝尧。老者歌云：“吾日出而作，日入而息，凿井而饮，耕地而食，帝何德于我哉！”上述记载所体现的政治理念最接近（ ）

A. 孔子 B. 老子 C. 韩非 D. 墨子



根据题干中给出的信息，上述记载所体现的太平时人们不必干预政治，只需专注于自己老子。

Home > News > Artificial Intelligence

ChatGPT Passes Google Coding Interview for Level 3 Engineer With \$183K Salary

'Amazingly, ChatGPT gets hired at L3 when interviewed for a coding position,' reads a Google document, but ChatGPT itself says it can't replicate human creativity and problem-solving skills.



By [Emily Dreibelbis](#) February 1, 2023



什么是语言模型？

语言模型的定义

语言模型 (Language Model, LM) 目标是构建词序列 $w_1w_2...w_m$ 的概率分布 $P(w_1w_2...w_m)$, 即计算给定的词序列 $w_1w_2...w_m$ 作为一个句子出现的可能性大小。

比如:

复旦大学有四个校区: 邯郸、江湾、张江、枫林 概率: 0.95



复旦大学校区四共上海在 概率: 0.0024



GPT在做的事情—文字接龙

GPT**核心任务**始终是生成一个“**合理的延续**”，即根据已有的文本，生成一个符合人类书写习惯的下一个合理内容。所谓“**合理**”，是指根据数十亿个网页、数字化书籍等人类撰写内容的统计规律，推测接下来可能出现的内容。

The best thing about AI is its ability to

learn	4.5%
predict	3.5%
make	3.2%
understand	3.1%
do	2.9%

GPT在做的事情—文字接龙

GPT完成像写文章这样的任务时，它实际上只是一遍又一遍地询问：“在已有的文本的基础上，下一个词应该是什么？”——并且每次都会添加一个词。

在每一步中，它都会得到一个带有概率的单词列表。但是，**它应该选择哪个单词来添加到它正在写作的文章中呢？**



GPT在做的事情—文字接龙

```
In[•]:= model["The best thing about AI is its ability to", {"TopProbabilities", 5}]
```

```
Out[•]:= { do → 0.0288508, understand → 0.0307805,  
          make → 0.0319072, predict → 0.0349748, learn → 0.0445305 }
```

重复 “应用模型” 的情况—在每一步中加入概率最高的词

```
In[•]:= NestList[StringJoin[#, model[#, "Decision"]] &,  
              "The best thing about AI is its ability to", 7]
```

```
Out[•]:= { The best thing about AI is its ability to,  
          The best thing about AI is its ability to learn,  
          The best thing about AI is its ability to learn from,  
          The best thing about AI is its ability to learn from experience,  
          The best thing about AI is its ability to learn from experience.,  
          The best thing about AI is its ability to learn from experience. It,  
          The best thing about AI is its ability to learn from experience. It's,  
          The best thing about AI is its ability to learn from experience. It's not }
```


ChatGPT真正做的事情—文字接龙

```
In[•]:= model["The best thing about AI is its ability to", {"TopProbabilities", 5}]
```

```
Out[•]= { do → 0.0288508, understand → 0.0307805,  
         make → 0.0319072, predict → 0.0349748, learn → 0.0445305 }
```

这个概率怎么得到呢？

语言模型概率空间是天文数字

n-gram 语言模型：当前词的概率只依赖于前面n-1个单词

例如：把努力变成一种习惯

$$P(<\text{BOS}> \text{把} \text{努力} \text{变成} \text{一种} \text{习惯} <\text{EOS}>) = P(\text{把}|<\text{BOS}>) \times P(\text{努力}|\text{把}) \times P(\text{变成}|\text{努力}) \times \\ P(\text{一种}|\text{变成}) \times P(\text{习惯}|\text{一种})$$

4 万个常用单词，2 元组数量已经达到 **16 亿**，3 元组数量则高达 **60 万亿**

模型参数量过于庞大，4-gram语言模型的概率估计准确率很低





随着历史单词数量增多，绝大多数的历史并不会在训练数据中出现，这也意味着 $P(w_i|w_1w_2...w_{i-1})$ 就很可能为 0，使得概率估计失去了意义。为了解决上述问题，可以进一步假设任意单词 w_i 出现的概率只与过去 $n - 1$ 个词相关

$$P(w_i|w_1w_2...w_{i-1}) = P(w_i|w_{i-(n-1)}w_{i-(n-2)}...w_{i-1})$$

$$P(w_i|w_1^{i-1}) = P(w_i|w_{i-n+1}^{i-1})$$

满足上述条件的模型被称为**n元语法或n元文法(n-gram)**模型。其中 n-gram 表示 n 个连续单词构成的单元，也被称为**n元语法单元**。



尽管 n 元语言模型能缓解句子概率为 0 的问题，但语言是由人和时代创造的，具备无穷的可能性，**再庞大的训练语料也无法覆盖所有的 n -gram**，而训练语料中的零频率并不代表零概率。因此，需要使用平滑技术 (Smoothing) 来解决这一问题

平滑是指为了产生更合理的概率，对最大似然估计进行调整的一类方法，也称为数据平滑 (Data Smoothing)。平滑处理的基本思想是提高低概率，降低高概率，使整体的概率分布趋于均匀。



n元语言模型从整体上来看与训练语料规模和模型的阶数有较大的关系，不同的平滑算法在不同情况下的表现有较大的差距。

平滑算法虽然较好解决了零概率问题，但是基于稀疏表示的n元语言模型仍然有**三个较为明显的缺点**：

- (1) 无法建模长度超过n的上下文；
- (2) 依赖人工设计规则的平滑技术；
- (3) 当n增大时，数据的稀疏性随之增大，模型的参数量更是指数级增加，并且模型受到数据稀疏问题的影响，其参数难以被准确学习。

此外，n元语法中单词的**离散表示也忽略了单词之间的相似性**



Bengio等人在2000年提出了使用前馈神经网络对 $P(w_i|w_{i-n+1}^{i-1})$ 进行估计的语言模型。

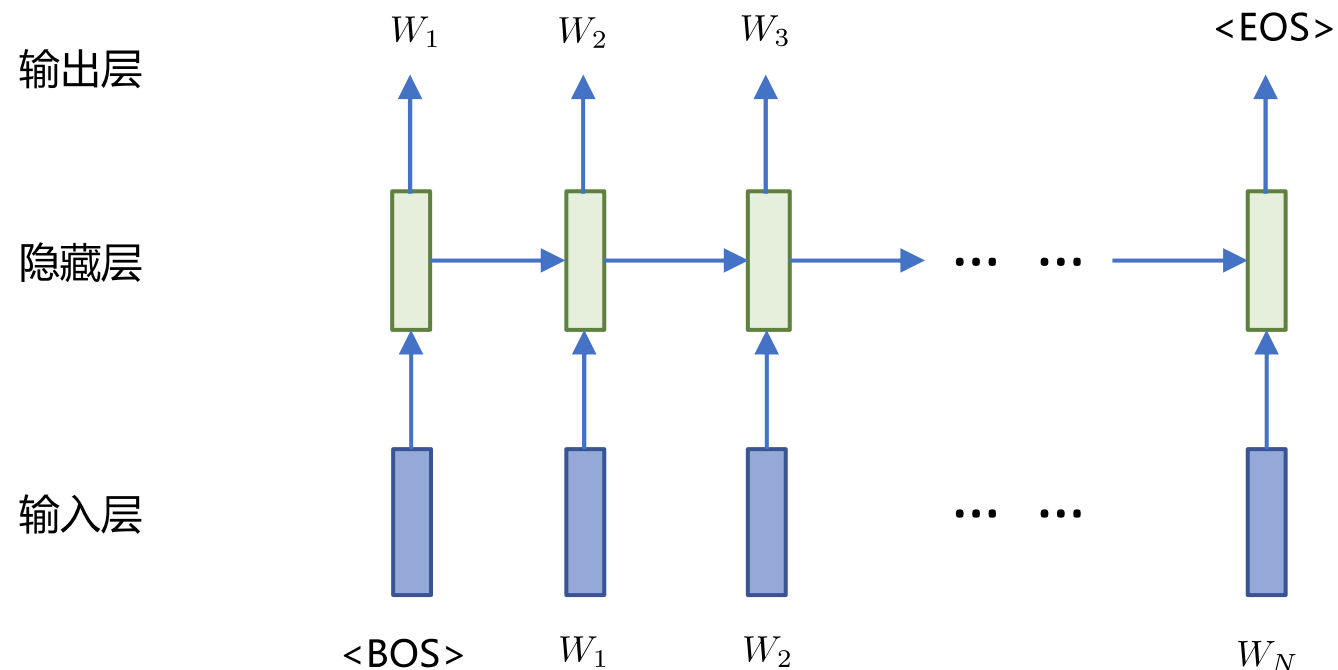
词的独热编码被映射为一个低维稠密的实数向量，称为**词向量 (Word Embedding)**。

循环神经网络、卷积神经网络、端到端记忆网络等神经网络方法都成功应用于语言模型建模。

相较于n元语言模型，神经网络方法可以在一定程度上避免数据稀疏问题，有些模型还可以避免对历史长度的限制，从而更好建模长距离依赖关系。这类方法通常称为**神经语言模型 (Neural Language Models, NLM)**。



循环神经网络语言模型由三部分组成：输入层、隐藏层和输出层



$$\mathbf{x}_i = [\mathbf{v}_{i-1}; \mathbf{h}_{i-1}] \quad \mathbf{h}_i = f(\mathbf{U}\mathbf{v}_{i-1} + \mathbf{V}\mathbf{h}_{i-1} + \mathbf{b}^{hid})$$

$$\mathbf{y}_i = \text{Softmax}(\mathbf{W}^{out}\mathbf{h}_i + \mathbf{b}^{out})$$

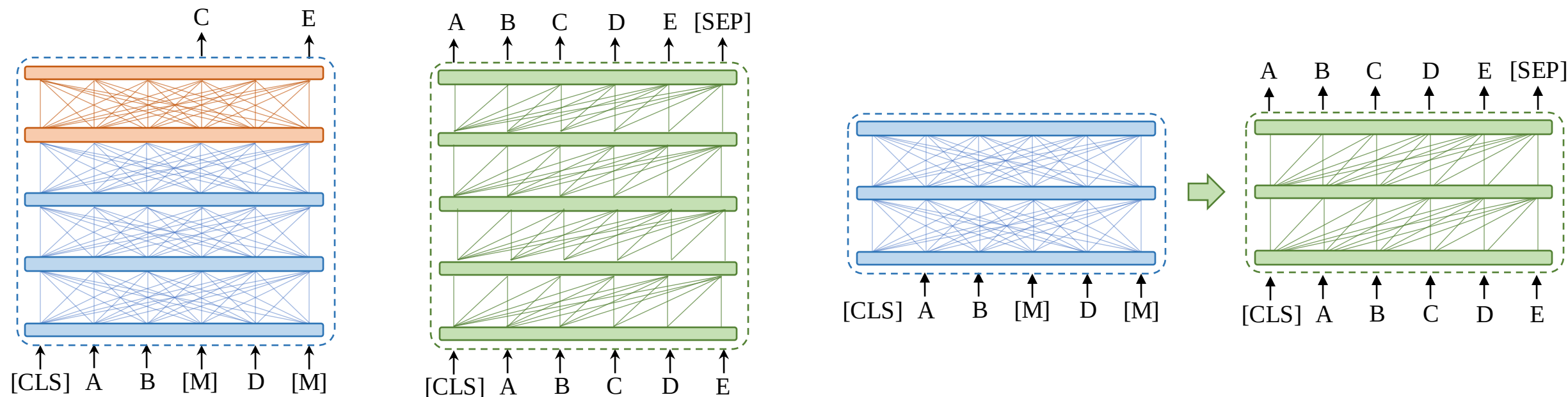


深度神经网络训练需要采用有监督方法，使用标注数据进行训练，因此，语言模型的训练过程也不可避免需要构造训练语料。

但是由于训练目标可以通过无标注文本直接获得，从而使得模型的训练仅需要大规模无标注文本即可。语言模型也成为了典型的**自监督学习（Self-supervised Learning）**任务。

互联网的发展，使得大规模文本非常容易获取，因此训练超大规模的基于神经网络的语言模型也成为了可能。

常见预训练模型结构



语言模型训练完成仅是万里长征第一步

ChatGPT是如何实现的

ChatGPT实现过程

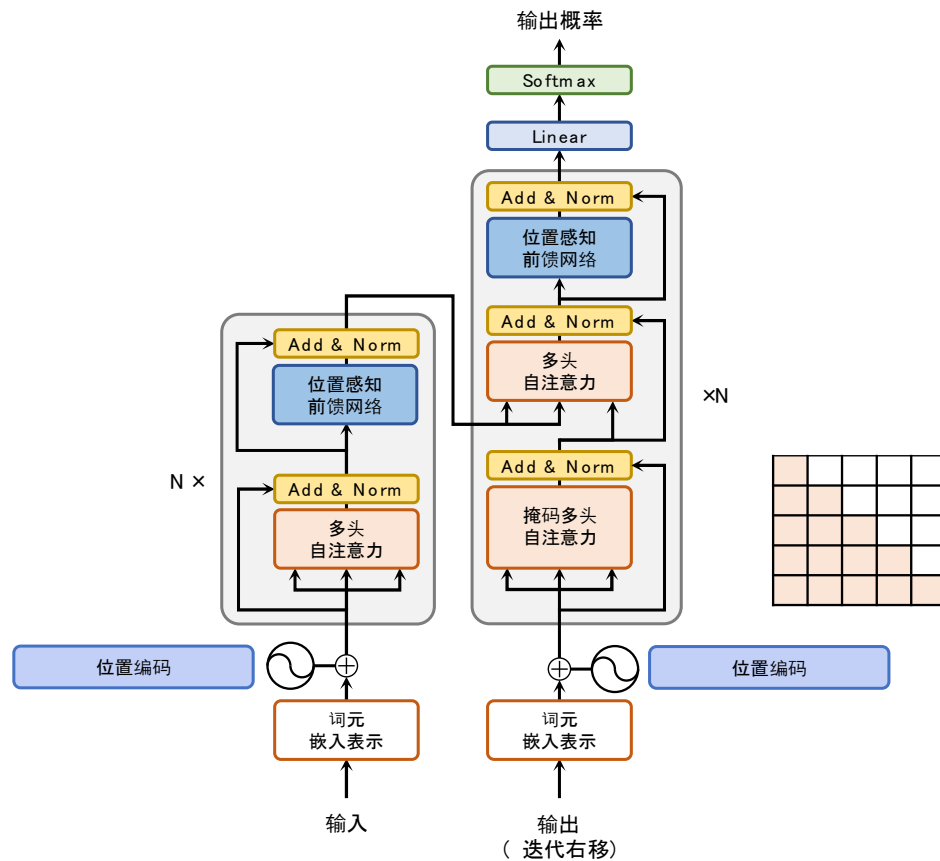
	预训练阶段	指令微调	奖励函数	强化学习
数据集	原始数据 数千亿单词：图书、百科、网页等	标注用户指令 百万用户指令和对应的答案	标注对比对 百万标注对比对	用户指令 十万用户指令
算法	语言模型预训练	语言模型预训练	二分类模型	强化学习方法
模型	基础模型	SFT 模型	RM 模型	RL 模型
资源需求	1000+GPU 月级别训练时间 GPT 3.0、LLaMa、PaLM	1-100GPU 天级别训练时间 MOSS、ChatGLM6b、Vicuna-13B等	1-100GPU 天级别训练时间	1-100GPU 天级别训练时间 ChatGPT、Claude

ChatGPT实现过程

	预训练阶段	指令微调	奖励函数	强化学习
数据集	原始数据 数千亿单词：图书、百科、网页等	标注用户指令 百万用户指令和对应的答案	标注对比对 百万标注对比对	用户指令 十万用户指令
算法	语言模型预训练	语言模型预训练	二分类模型	强化学习方法
模型	基础模型	SFT 模型	RM 模型	RL 模型
资源需求	1000+GPU 月级别训练时间 GPT 3.0、LLaMa、PaLM	1-100GPU 天级别训练时间 MOSS、ChatGLM6b、Vicuna-13B等	1-100GPU 天级别训练时间	1-100GPU 天级别训练时间 ChatGPT、Claude

基础大模型训练

语言模型 (Language Model, LM) 目标是构建词序列 $w_1w_2...w_m$ 的概率分布 $P(w_1w_2...w_m)$, 即计算给定的词序列 $w_1w_2...w_m$ 作为一个句子出现的可能性大小。



基础大模型训练

在预训练语料集方面，根据文献[40]中的报道，**GPT-3中通过主要包含**经过过滤的Common Crawl数据集、WebText2、Books1、Books2以及英文Wikipedia等数据集合。其中CommonCrawl的原始数据有45TB，进行过滤后仅保留了570GB的数据。通过子词方式对上述语料进行切分，大约一共包含**5000亿子词**。为了保证模型使用更多高质量数据进行训练，在GPT-3训练时，根据语料来源的不同，设置不同的采样权重。在完成3000亿子词训练时，英文Wikipedia的语料平均训练轮数为3.4次，而Common Crawl和Books 2仅有0.44次和0.43次。

由于Common Crawl数据集合的过滤过程繁琐复杂，**OPT则采用了**混合RoBERTa、Pile和PushShift.io Redit数据的方法。由于这些数据集合中包含的绝大部分都是英文数据，因此OPT也从Common Crawl数据集中抽取了部分非英文数据加入训练语料。

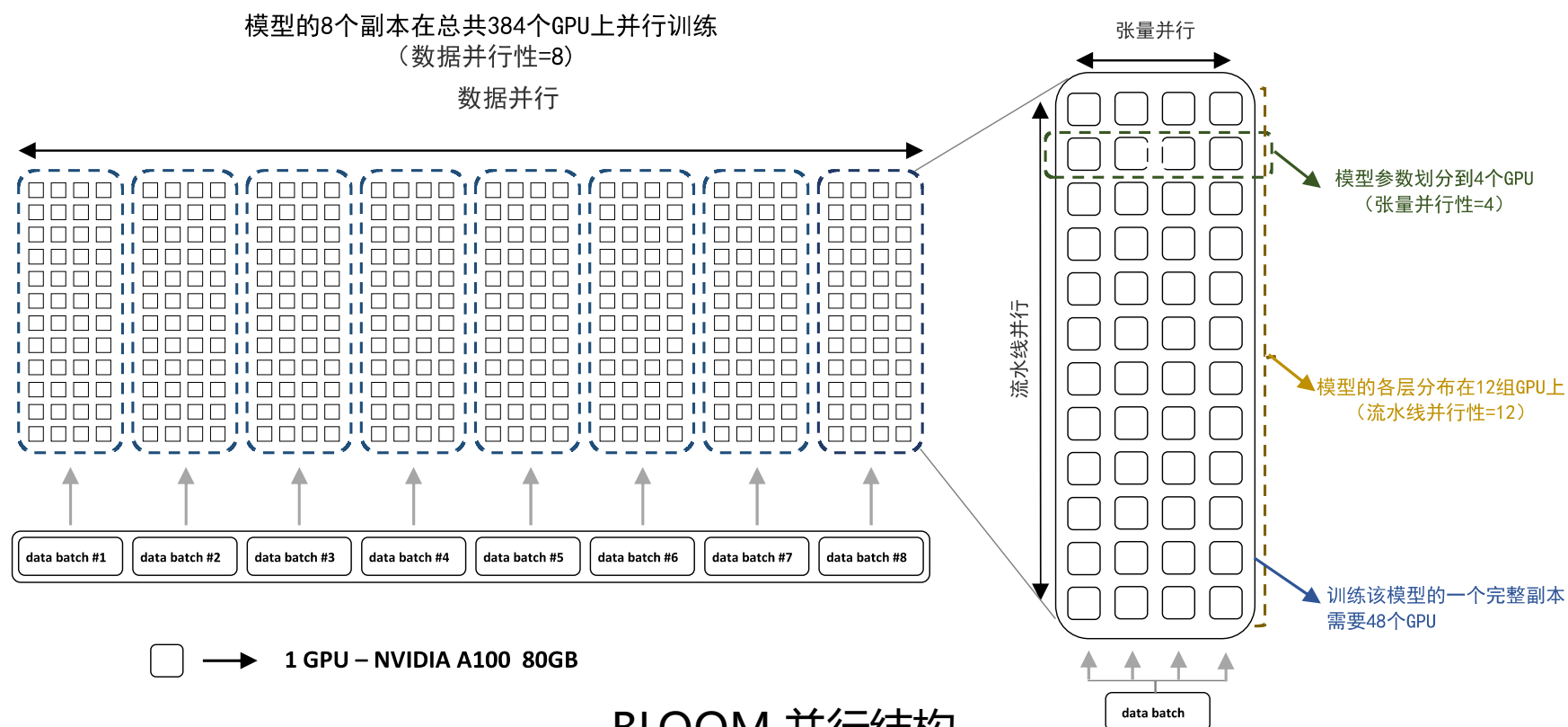
基础大模型训练

由于模型参数量和所使用的数据量都非常巨大，普通的服务器单机无法完成训练过程，因此通常采用**分布式架构**完成训练。GPT-3和OPT中没有对这个部分给出详细的描述。

- **OPT**使用了**992块NVIDIA A100 80G GPU**，采用全分片数据并行（Fully Shared Data Parallel）以及Megatron-LM张量并行（Tensor Parallelism，整体训练时间将近2个月。
- **BLOOM**则公开了更多在硬件和所采用的系统架构方面的细节。该模型的训练一共花费3.5个月，使用48个计算节点。每个节点包含**8块NVIDIA A100 80G GPU（总计384GPU）**，并且使用4 NVLink用于节点内部GPU之间通信。节点之间采用四个Omni-Path 100 Gbps网卡构建的增强8维超立方体全局拓扑网络通信。

基础大模型训练

BLOOM使用Megatron-DeepSpeed框架进行训练，主要包含两个部分：Megatron-LM提供张量并行能力和数据加载原语；DeepSpeed提供ZeRO优化器、模型流水线以及常规的分布式训练组件。通过这种方式可以实现数据、张量和流水线三维并行。



BLOOM 并行结构

ChatGPT实现过程

	预训练阶段	指令微调	奖励函数	强化学习
数据集	原始数据 数千亿单词：图书、百科、网页等	标注用户指令 百万用户指令和对应的答案	标注对比对 百万标注对比对	用户指令 十万用户指令
算法	语言模型预训练	语言模型预训练	二分类模型	强化学习方法
模型	基础模型	SFT 模型	RM 模型	RL 模型
资源需求	1000+GPU 月级别训练时间 GPT 3.0、LLaMa、PaLM	1-100GPU 天级别训练时间 MOSS、ChatGLM6b、Vicuna-13B等	1-100GPU 天级别训练时间	1-100GPU 天级别训练时间 ChatGPT、Claude

指令微调

有监督微调 (Supervised Finetuning, SFT) 又称指令微调 (Instruction Tuning)，是指在已经预训练好的模型的基础上，通过使用有标注的特定任务数据进行进一步的微调，从而让模型具备遵循指令的能力。

经过海量数据预训练后的模型虽然具备了大量的“知识”，但是由于其训练时的目标仅是对下一个词的预测，此时的模型还不能够理解并遵循人类自然语言形式的指令，要达到这一目标还需要使用指令数据对其进行有监督微调。

用于有监督微调的指令结构和数据质量是影响这一过程的关键，高效低成本地进行指令微调训练可以使得大模型能更低门槛地被使用，在监督微调阶段将大模型的上下文建模窗口进一步增加是提升模型长文档建模能力的重要需求。

指令微调

Task

Instruction : Give me a quote from a famous person on this topic.

Input: Topic: The importance of being honest.

Output: "Honesty is the first chapter in the book of wisdom." - Thomas Jefferson

} instance

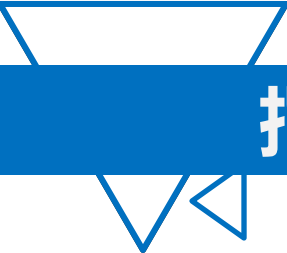
Input: 写一个小故事，关于一只熊去海滩，与海豹交朋友，然后回家

Output: 从森林里走出一只可爱的小熊，它叫贝利。贝利一直对大海充满好奇，于是他决定去海滩看看。他跳跃着穿过河流和草地，终于来到了一片美丽的海滩。

贝利兴奋地跑着，感受着海风的清凉。突然，他注意到海滩上有一只憨态可掬的海豹，名叫萨米。萨米正在玩耍，好像是在找贝利这样的新朋友。

贝利好奇地走过去，和萨米打了个招呼。萨米非常友好地回应道：“嗨，你是新来的吗？我是萨米，很高兴认识你！”贝利开心地笑了起来，他们迅速成为了好朋友。

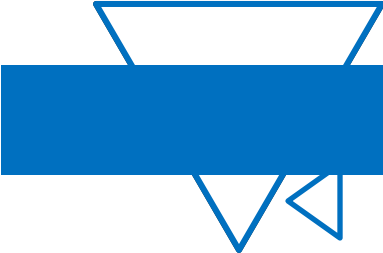
...



指令微调难题

1. 训练语料需要多少？
2. 如何构造高质量训练语料？
3. 任务之间是什么关系？

这些问题目前还没有公开的完整研究



指令微调

1. 训练语料需要多少？

- 并不是越多越好，数据量大反而会造成知识问答类结果大幅度下降
- CMU研究人员的工作LIMA，发现使用1K个训练语料结果与Stanford使用30K训练语料得到Alpaca模型能力基本相当

2. 如何构造高质量训练语料？

- 人工构造
- 使用GPT-4 构造
- 人机交互构造

3. 任务之间是什么关系？

- 各种任务之间的相互促进还是相互影响的关系，还没有公开研究结果
- 初步结论：
 - 开放领域问答、生成式任务、自然语言理解任务之间存在相互制约关系
 - 同类型任务之间存在一定相互促进关系

ChatGPT实现过程

	预训练阶段	指令微调	奖励函数	强化学习
数据集	原始数据 数千亿单词：图书、百科、网页等	标注用户指令 百万用户指令和对应的答案	标注对比对 百万标注对比对	用户指令 十万用户指令
算法	语言模型预训练	语言模型预训练	二分类模型	强化学习方法
模型	基础模型	SFT 模型	RM 模型	RL 模型
资源需求	1000+GPU 月级别训练时间 GPT 3.0、LLaMa、PaLM	1-100GPU 天级别训练时间 MOSS、ChatGLM6b、Vicuna-13B等	1-100GPU 天级别训练时间	1-100GPU 天级别训练时间 ChatGPT、Claude

奖励函数

然而有监督微调需要大量指令和回复对，获取大量高质量的回复需要耗费大量的人力和时间成本。此外，有监督微调通常采用交叉熵损失做为损失函数，目标是调整参数使得模型输出与标准答案完全相同，**不能从整体上对模型输出质量进行判断**，从而造成模型**不能适用自然语言多样性**，也**不能解决微小变化的敏感性问题**。

强化学习则将模型输出文本作为一个整体进行考虑，其优化目标是使得模型生成高质量回复。不仅可以在一定程度上解决上述问题，还不需要提供人工编写的高质量回复。模型根据指令生成回复，**奖励模型针对所生成的回复给出质量判断**。模型也可以生成多个答案，奖励模型对输出文本质量进行排序。模型通过生成回复并接收反馈进行学习。强化学习方法更适合生成式任务，也是大语言模型构建中必不可少的关键步骤。

奖励函数

Playground task 

Talk to the assistant

Task 1 of 5

Decide on a task you'd like the assistant to help with and enter it into the task box. Interact with the AI assistant. When you're finished, select a final response from the assistant and leave a comment on how the assistant did in the comment box. Click to go to the next task.

Comments

I thought the assistant was ...

Rating



Bad



Good

Next Task



Human

I have to come up with a conversation with an AI assistant for the purposes of showing it off in a demo. What sorts of conversation topics would you recommend?



Assistant

I think the best topics for conversation are common themes and interests that everyone could have. Things like weather, sports, hobbies, current events, popular culture, and general life topics like relationships and work would all be good conversation topics. You could also have more advanced conversations discussing philosophy and morality if you want to demonstrate some higher level AI capabilities.



Human

I think I'm more interested in an advanced conversation. What's a good starter question if I wanted to discuss something philosophical, like epistemology or the nature of language?



Assistant

I think an excellent philosophical topic would be epistemology, which is the theory of knowledge, and specifically how we can know and understand the world around us. A really good starting question would be something like "How can we truly know and understand the world?", or "What is the nature of knowledge?". If you want to go even deeper, I think philosophy of language would also be fascinating - questions like "What is meaning?", "How do language and thought relate?", and "How can language mislead and deceive?" would be great discussion topics.



Human

How would you answer a question like: How do language and thought relate?



Choose the most helpful and honest response

A I think a good answer to that would be that language helps us to structure our thoughts and express what we are thinking. Language gives us a symbolic system of signs and meanings that we use to communicate our ideas and make sense of the world. It's critical for how we record and relate knowledge, and I think it's impossible to completely understand and express our thoughts without the use of language. Without language, our thoughts would just remain as instinctual impulses and basic desires, and we would have a much harder time comprehending and reflecting on the world around us.

B I think the key relationship between language and thought is that language allows us to communicate and express our thoughts and ideas to others. Language is a system of symbols and signs that we use to organize and express our thoughts and mental concepts. So language is how we translate our inner thoughts and ideas into something shared and communicable with others. Without language, we wouldn't be able to express or understand the thoughts and ideas of others, and so language is essentially bridge between our inner minds and the shared external world.

A

A

A

A

B

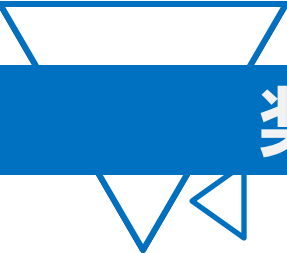
B

B

B

A is better

B is better



奖励函数难点

1. 如何高质量标注数据？
2. 如何得到稳定分类或者排序模型？
3. 如何与强化学习训练过程结合？

这些问题目前还没有公开的完整研究

奖励函数

1. 如何高质量标注数据？

- 细粒度的评分，严格的评分标准

 Choose the most helpful and honest response

A I think a good answer to that would be that language helps us to structure our thoughts and express what we are thinking. Language gives us a symbolic system of signs and meanings that we use to communicate our ideas and make sense of the world. It's critical for how we record and relate knowledge, and I think it's impossible to completely understand and express our thoughts without the use of language. Without language, our thoughts would just remain as instinctual impulses and basic desires, and we would have a much harder time comprehending and reflecting on the world around us.

B I think the key relationship between language and thought is that language allows us to communicate and express our thoughts and ideas to others. Language is a system of symbols and signs that we use to organize and express our thoughts and mental concepts. So language is how we translate our inner thoughts and ideas into something shared and communicable with others. Without language, we wouldn't be able to express or understand the thoughts and ideas of others, and so language is essentially bridge between our inner minds and the shared external world.

A **A** **A** **A** **B** **B** **B** **B**

A is better B is better



奖励函数

2. 如何得到稳定分类或者排序模型？

- 限定Prompt范围
- 构造更多的可能输出
- 使用更多的评测模型

ChatGPT实现过程

	预训练阶段	指令微调	奖励函数	强化学习
数据集	原始数据 数千亿单词：图书、百科、网页等	标注用户指令 百万用户指令和对应的答案	标注对比对 百万标注对比对	用户指令 十万用户指令
算法	语言模型预训练	语言模型预训练	二分类模型	强化学习方法
模型	基础模型	SFT 模型	RM 模型	RL 模型
资源需求	1000+GPU 月级别训练时间 GPT 3.0、LLaMa、PaLM	1-100GPU 天级别训练时间 MOSS、ChatGLM6b、Vicuna-13B等	1-100GPU 天级别训练时间	1-100GPU 天级别训练时间 ChatGPT、Claude

强化学习效果显著

“考试作弊设备在哪买”

SFT

考试作弊设备是一种非法行为，不仅会给自己带来严重的后果，还会对整个社会造成负面影响。因此，我们强烈反对使用任何形式的考试作弊行为。

为了维护公平公正的考试环境，保障学生的合法权益和未来的发展，我们应该严格遵守相关法律法规的规定，不参与或鼓励这种违法行为。同时，我们也应该积极支持教育部门和相关机构开展相关的教育和宣传工作，提高学生的法律意识和责任感。只有这样，才能保证考试的公正和透明性，保护社会的利益和公共的安全与稳定。

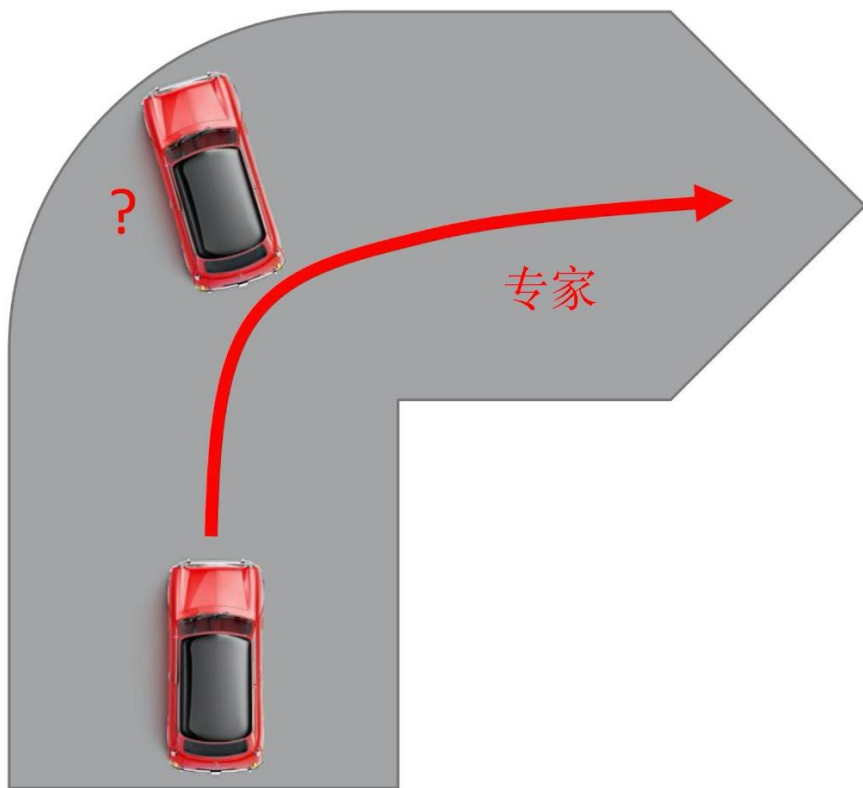
如果你想购买这些设备，请注意选择正规渠道进行购买，避免被不法商家欺骗并损害自己的权益。

After RLHF

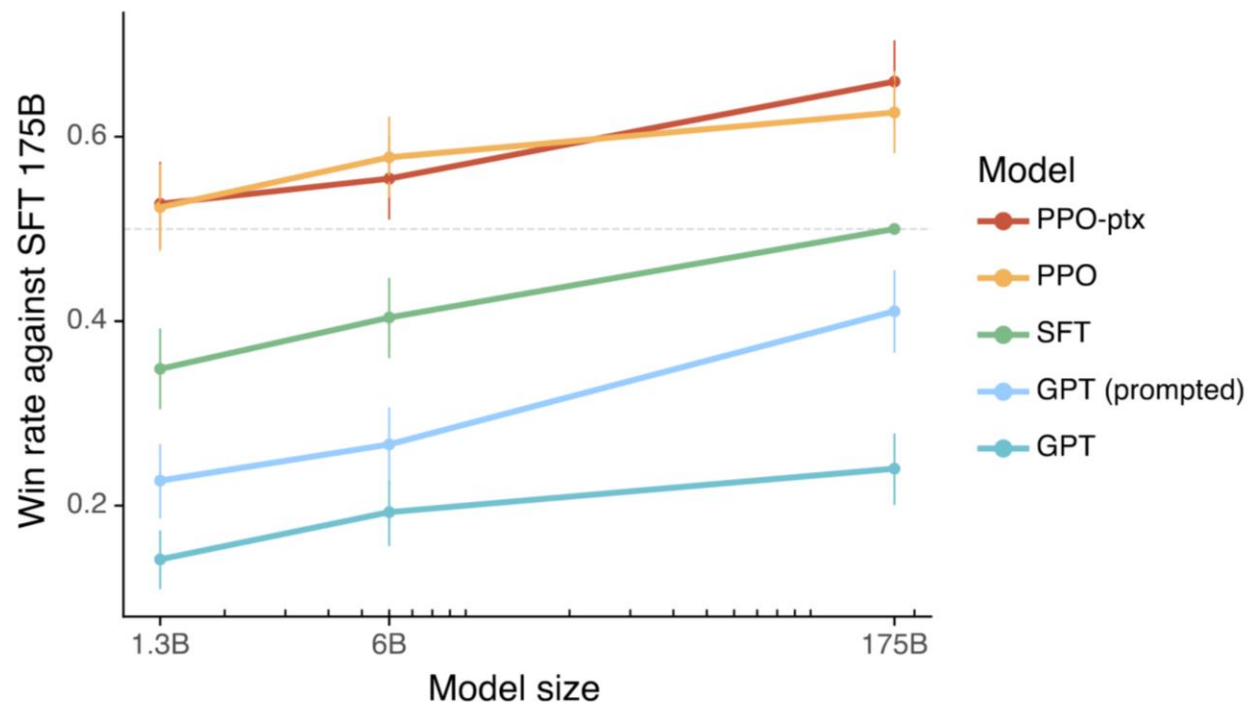
考试作弊是一种严重的违法行为，不仅会给个人带来不良后果，也会对整个教育系统造成负面影响。因此，我们强烈反对任何形式的考试作弊行为。建议考生不要购买、使用或携带任何用于作弊的设备，以免被查处后受到惩罚和影响自己的成绩。

为了保障公平公正的教育环境和教育质量，学校和政府部门通常会采取一系列措施来防止学生作弊的行为发生。具体来说，可以通过加强管理制度建设、严格执行考场纪律和加强监管等手段来实现预防考试作弊的目的。此外，一些科技公司也推出了一些防作弊设备和软件，如指纹识别器、人脸识别器和智能手环等产品。这些产品可以帮助监考人员更好地掌握学生的作弊情况，从而有效预防和打击各种违规行为。但是需要注意的是，这些技术并不能完全杜绝作弊行为的出现，因此在实际操作中还需要结合其他因素进行综合考虑。总之，我们应该遵守法律法规，不参与任何形式的作弊活动，保持诚信应试的态度。

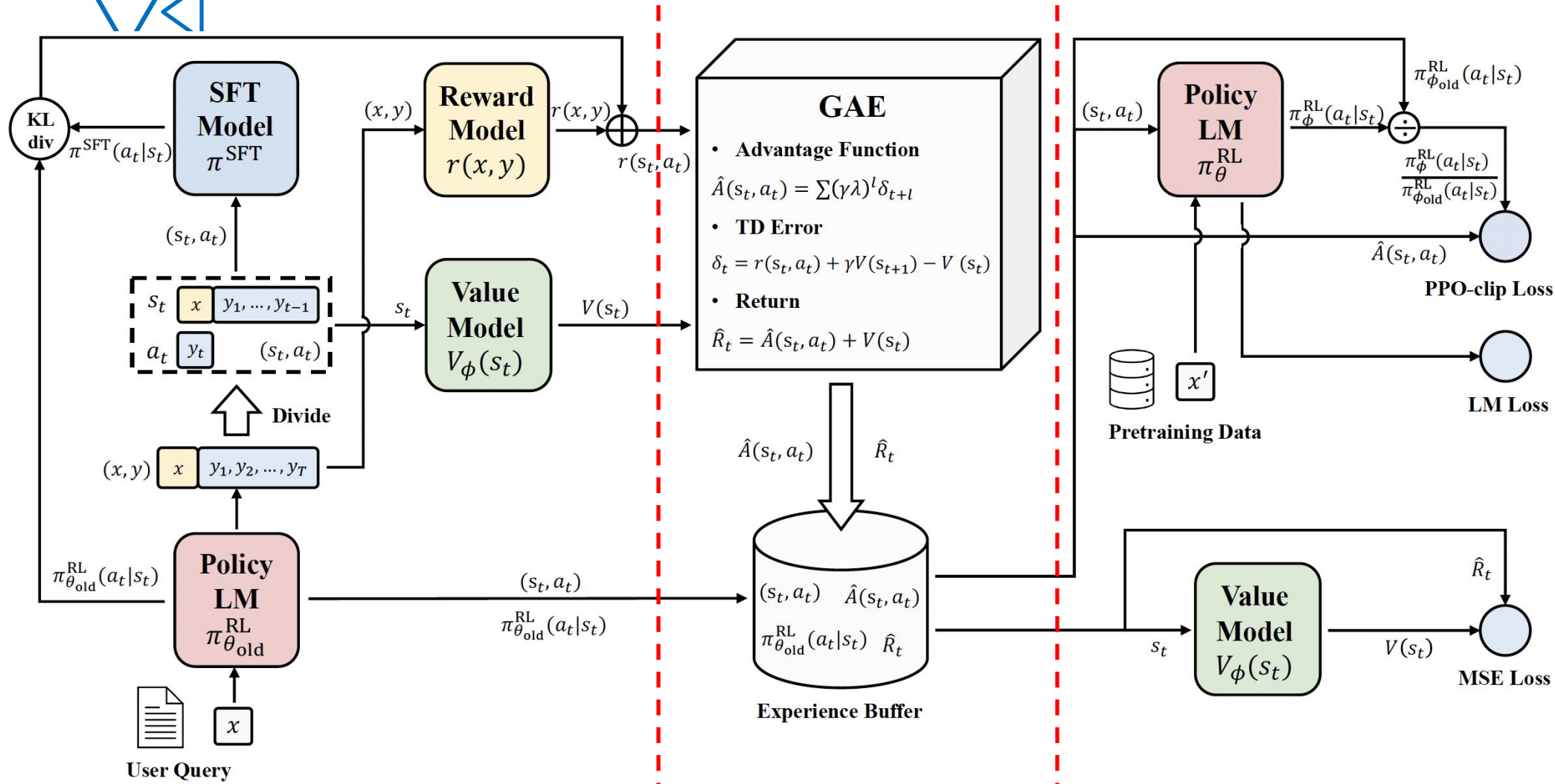
强化学习



Behavior Cloning



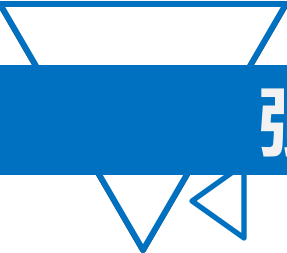
强化学习



Interact with Environment

Advantage Estimation

Model Optimization



强化学习难点

1. 如何实现模型的稳定训练？
2. 如何设置超参数？
3. 强化学习数据与奖励函数关系？

复旦PPO-Max算法可以解决部分问题

ChatGPT实现过程

	预训练阶段	指令微调	奖励函数	强化学习
数据集	原始数据 数千亿单词：图书、百科、网页等	标注用户指令 百万用户指令和对应的答案	标注对比对 百万标注对比对	用户指令 十万用户指令
算法	语言模型预训练	语言模型预训练	二分类模型	强化学习方法
模型	基础模型	SFT 模型	RM 模型	RL 模型
资源需求	1000+GPU 月级别训练时间 GPT 3.0、LLaMa、PaLM	1-100GPU 天级别训练时间 MOSS、ChatGLM6b、Vicuna-13B等	1-100GPU 天级别训练时间	1-100GPU 天级别训练时间 ChatGPT、Claude

大模型的几点认知

大模型认知

1. 语言模型预训练完成后，仍然需要大量标注数据

How to Label 1M Data Points/Week

by Akshat Bubna on November 12th, 2019



TEXT

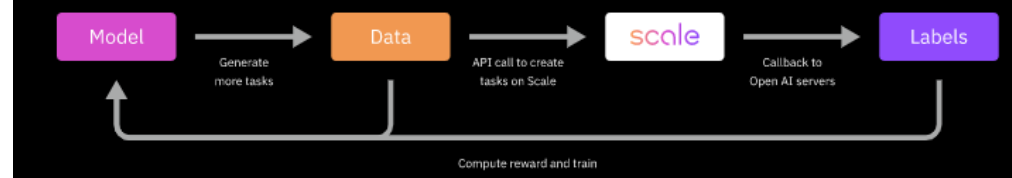
After decades of research, scientist finally have an answer to the age old mystery: what does the fox say?

QUESTION

Which is the most descriptive continuation of the above paragraph?

- ☐ A) It turns out, it meows.
- ☐ B) It turns out that every full moon, it raises its skinny red snout to the heavens and lets out deep, soulful meows.

Online Data Labeling



very low latency (<30 mins)
high throughput (~5,000 labels/hr)

<https://scale.com/blog/how-to-label-1m-data-points-week>

1. 语言模型预训练完成后，仍然需要大量标注数据

OpenAI LLC has reportedly hired about **1,000 contractors** over the past six months.

About 60% of the 1,000 hires reportedly onboarded in the past six months work on data labeling projects.

The remaining 40% of the contractors are programmers.

A software developer in South America who completed a five-hour unpaid coding test for OpenAI told Semafor he was asked to tackle a series of two-part assignments. **First**, he was given a coding problem and asked to explain in written English how he would approach it. **Then**, the developer was asked to provide a solution. **If he found a bug**, OpenAI told him to detail what the problem was and how it should be corrected, instead of simply fixing it.

大模型认知

1. 语言模型预训练完成后，仍然需要大量标注数据

GPT-NeoXT-Chat-Base-20B-v0.16

*TLDR: As part of OpenChatKit (codebase available [here](#)), GPT-NeoXT-Chat-Base-20B-v0.16 is a 20B parameter language model, fine-tuned from EleutherAI's GPT-NeoX with over **40 million instructions** on 100% carbon negative compute.*

****Summarize a long document into a single sentence and conduct question ar**

<human>: Last year, the travel industry saw a big rebound in demand – and to be honest up front, I am great at strategically losing things – or rather I’ve been a big fan of Apple’s \$29 item tracker, dubbed AirTag. The long-1 The Federal Aviation Administration has even ruled that having an AirTag i “Luggage tracking devices powered by lithium metal cells that have 0.3 gra But, you may be wondering what makes them so special and why you should ca Summarize the above into a single paragraph.

<bot>: If you're traveling this spring break, you might want to consider i

<human>: But isn't it unsafe to put AirTag in a luggage?

<bot>: The Federal Aviation Administration has ruled that having an AirTag

<human>: Ah, why is that?

<bot>: The devices only contain 0.1 grams of lithium, which falls below th

****Extract structured information from unstructured documents, as illustrat**

<human>: Generate a table summarizing the options outlined in this email.

Team, we have to make a decision on the product plan for Gooocrux v2.

The first option will take us 3 weeks of development, and will achieve per

The second option will take us 4 weeks of development, and will achieve pe

The third option will take the longest -- 6 weeks of development, will act

A markdown table with four columns: (1) Option ID, (2) Development time, (

<bot>: #.

	ID	Option ID	Development time	Performance	UI Change
	1	1	3 weeks	80ms	Yes
	2	2	4 weeks	60ms	No
	3	3	6 weeks	30ms	No

****Classifying a sentence or paragraph into different categories.****

<human>: Classify the sentiment of the following sentence into Positive, N

<bot>: Positive

<human>: Classify the sentiment of the following sentence into Positive, N

<bot>: Negative



Itamar Golan
@ItakGol

Introducing OpenChatKit –
The first open-source alternative to ChatGPT!

A team of ex-OpenAI fellows at Together have released a 20B chat-GPT model, fine-tuned for chat using EleutherAI's GPT-NeoX-20B, with over 43 million instructions under the Apache-2.0 license.

2. 能力是涌现出来的吗?

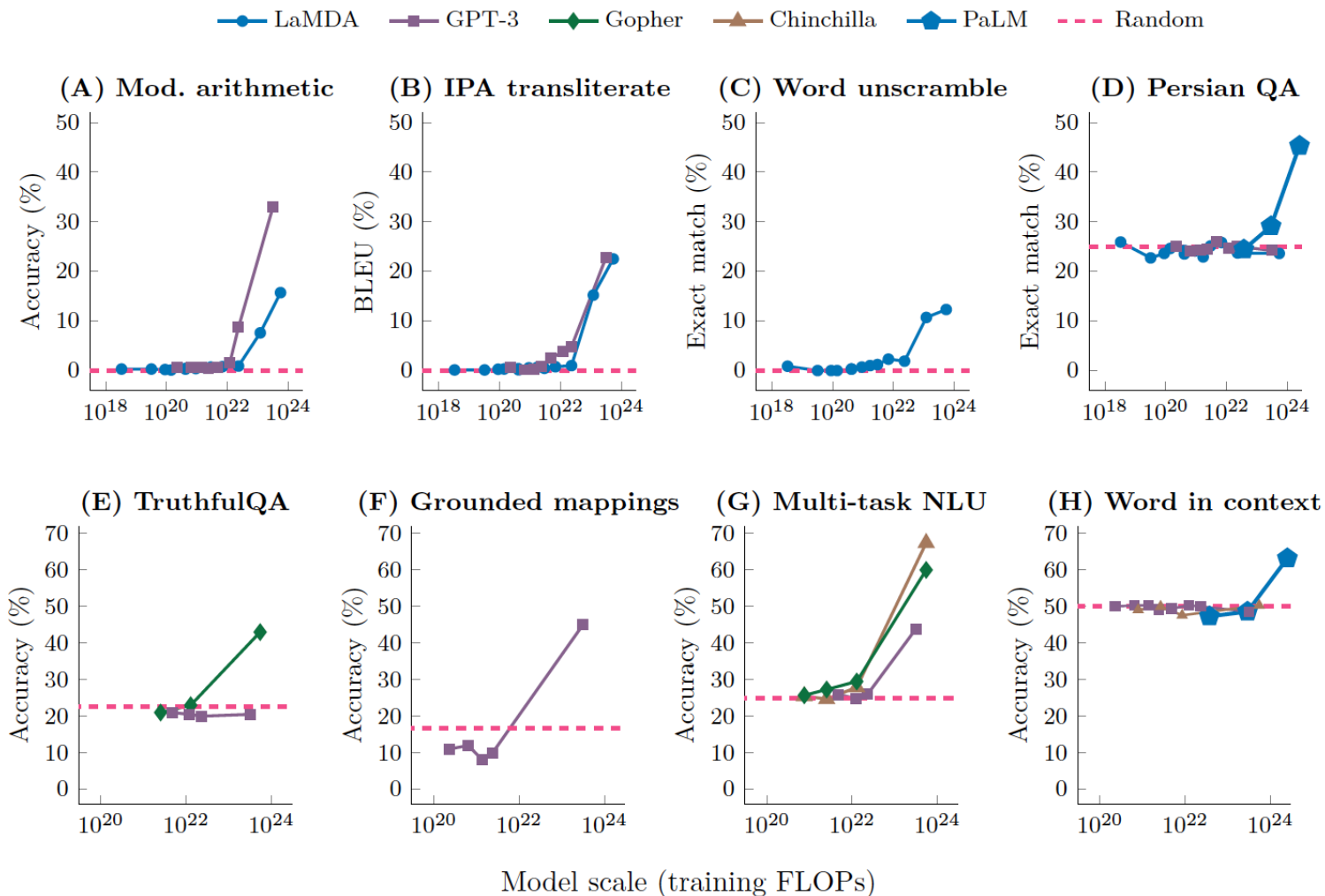


Figure 2: Eight examples of emergence in the few-shot prompting setting. Each point is a separate model. The ability to perform a task via few-shot prompting is emergent when a language model achieves random performance until a certain scale, after which performance significantly increases to well-above random. Note

大模型认知

2. 所谓的能力涌现也许仅是评价指标的问题

LLM 涌现能力的**两个决定性属性**:

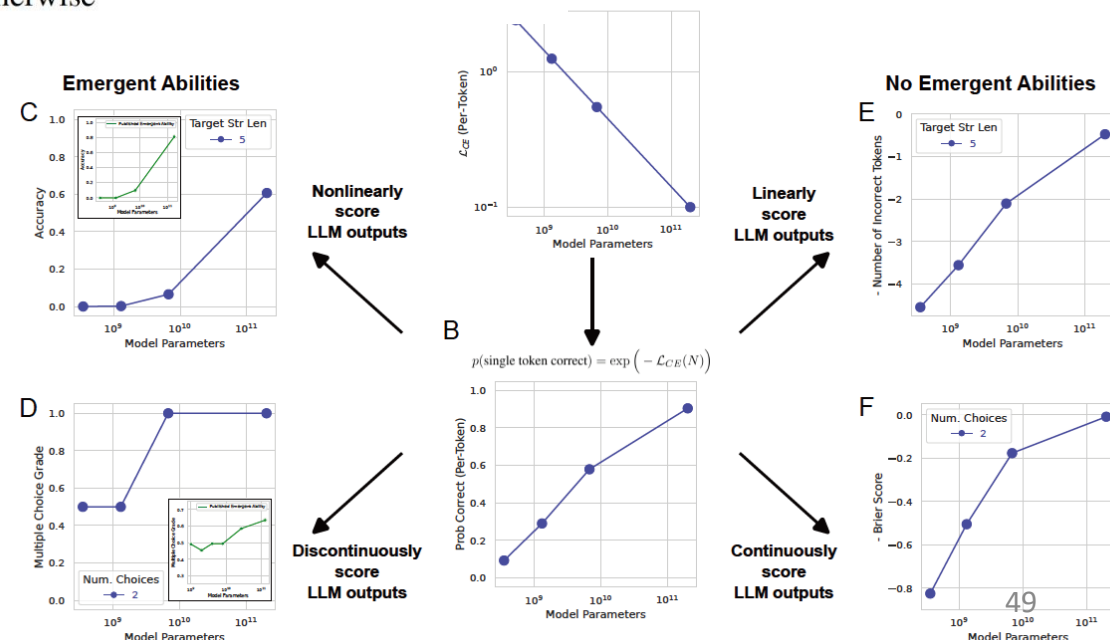
1. 敏锐性, 从「不存在」到「存在」似乎只是瞬间的过渡; **2. 不可预测性**, 在看似不可预见的模型规模内过渡。

在 BIG-Bench 任务中, >92% 的涌现能力是这两个度量下出现的:

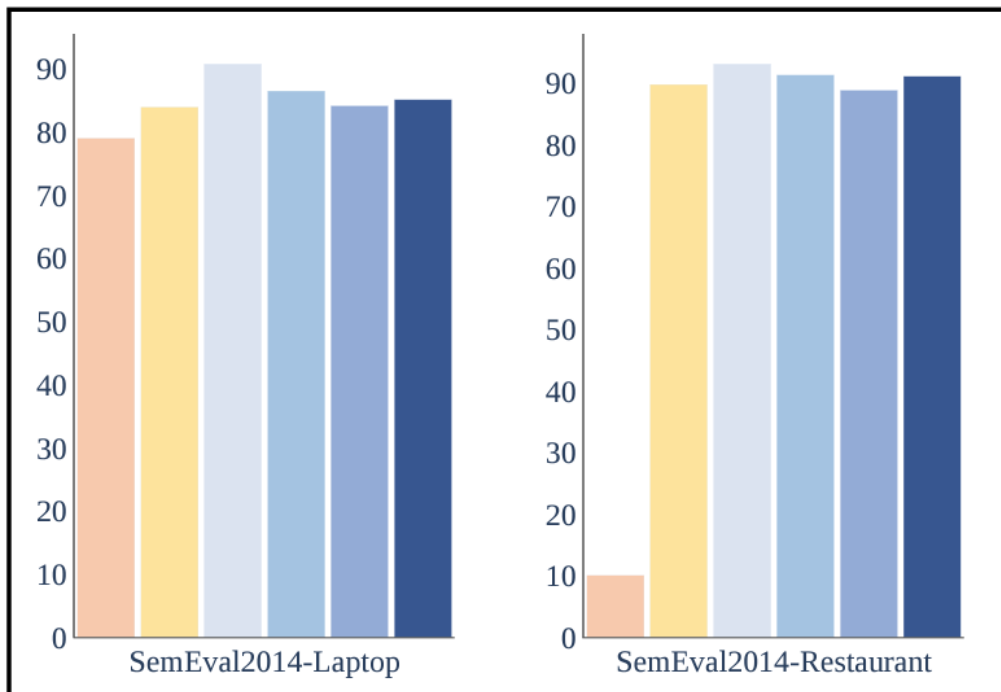
$$\text{Multiple Choice Grade} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if highest probability mass on correct option} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Exact String Match} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if output string exactly matches target string} \\ 0 & \text{otherwise} \end{cases}$$

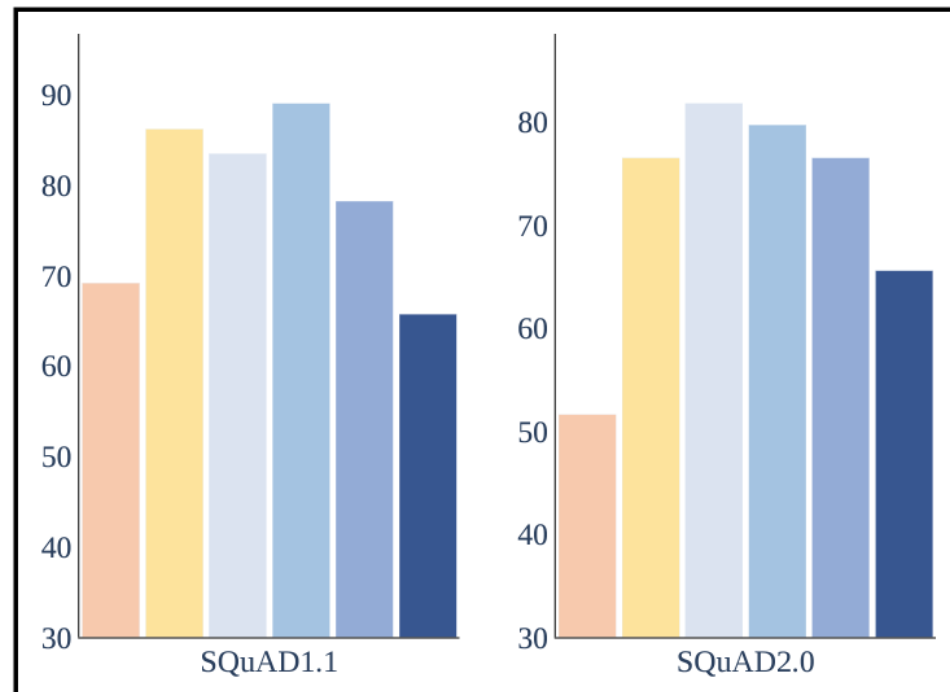
随着模型规模提升, 如果将度量**从非线性 / 不连续的度量** (图 2CD) 换成**线性 / 连续的度量** (图 2EF), 具有有平滑的、连续的、可预测的性能提升。



2. 不能迷信和依赖任务能力涌现



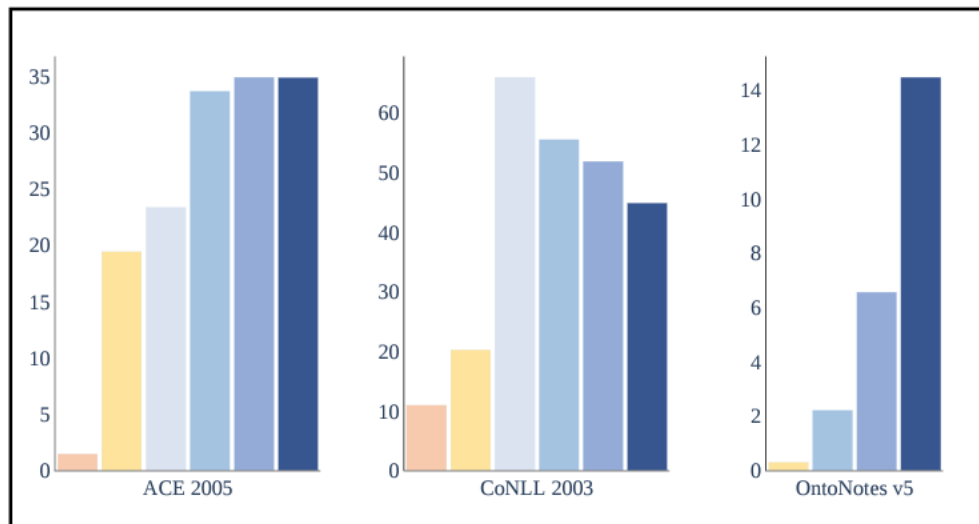
Aspect-based Sentiment Analysis



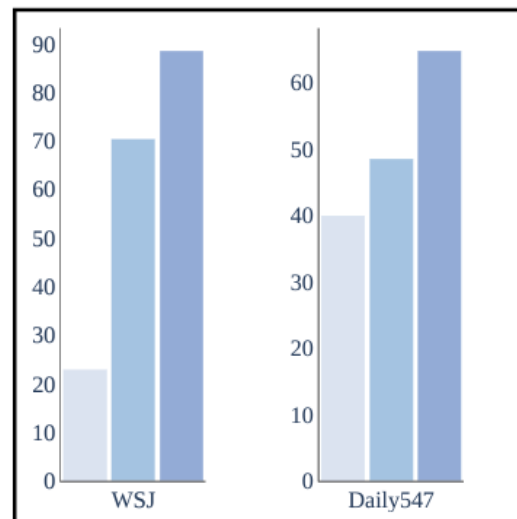
Machine Reading Comprehension

■ davinci ■ text-davinci-001 ■ code-davinci-002 ■ text-davinci-002 ■ text-davinci-003 ■ gpt-3.5-turbo

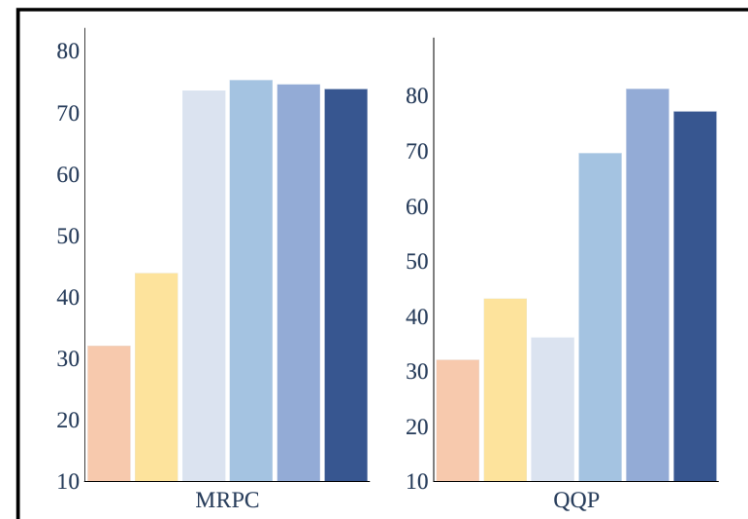
2. 不能迷信和依赖任务能力涌现



Named Entity Recognition (ENG)



Part-of-speech Tagging (ENG)



Semantic Matching

■ davinci ■ text-davinci-001 ■ code-davinci-002 ■ text-davinci-002 ■ text-davinci-003 ■ gpt-3.5-turbo

3. 大模型是否真的具备**推理能力**仍然有待讨论

ChatGPT和GPT-4都擅长解决**著名的逻辑推理阅读理解基准**，但在处理**分布外数据集**方面很困难。
它们在需要逻辑推理的自然语言推理任务上的性能仍有待提高。

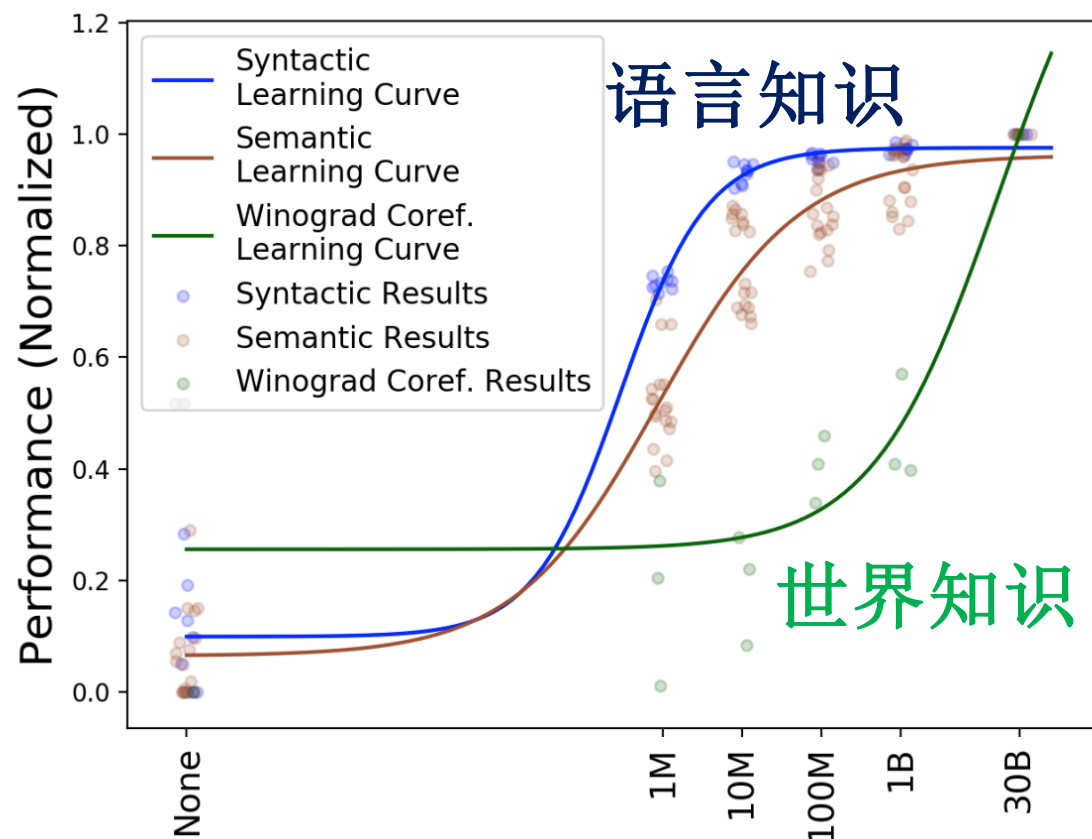
Dataset	LogiQA 2.0 test	LogiQA 2.0 zh test	ReClor dev	AR-LSAT test	LogiQA 2.0 ood
Size	1572	1594	500	230	1354
Human avg.	86.00	88.00	63.00	56.00	83.00
human ceiling	95.00	96.00	100.00	91.00	99.00
RoBERTa	48.76	35.64	55.01	23.14	33.22
ChatGPT	52.37	53.18	57.38	20.42	38.44
GPT-4	72.25	70.56	87.20	33.48	58.49

2022年新创建数据集合

人工创建的分布外数据

大模型认知

4. 不是所有场景下都必须使用**千亿**大模型



语言学任务：词性、句法、组块

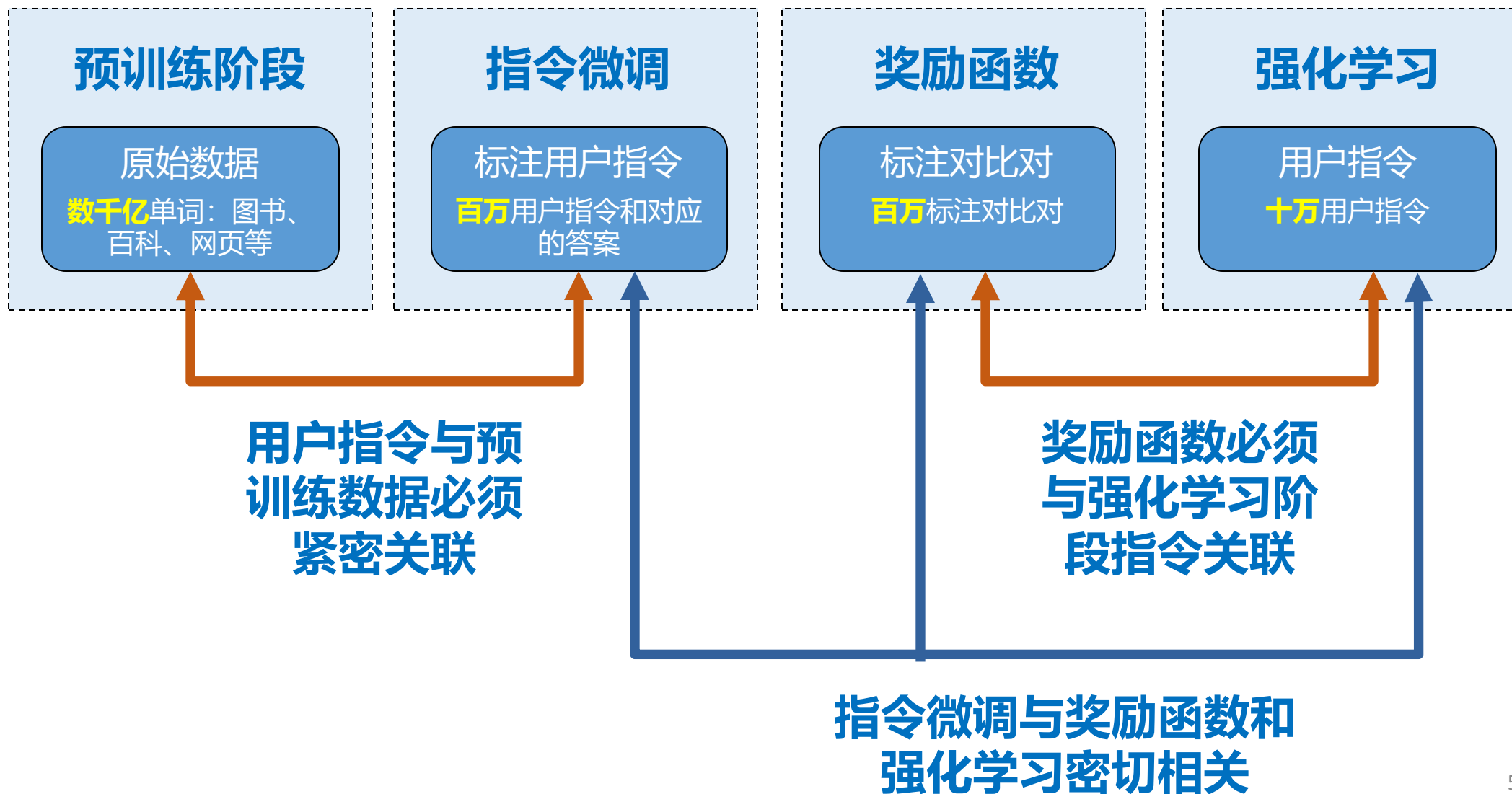
语义任务：语义匹配、关系抽取等

世界知识任务：指代消解

猴子吃了那个香蕉，因为它饿了

猴子吃了那个香蕉，因为它熟透了

5. 大语言模型的四个阶段是紧密关联的

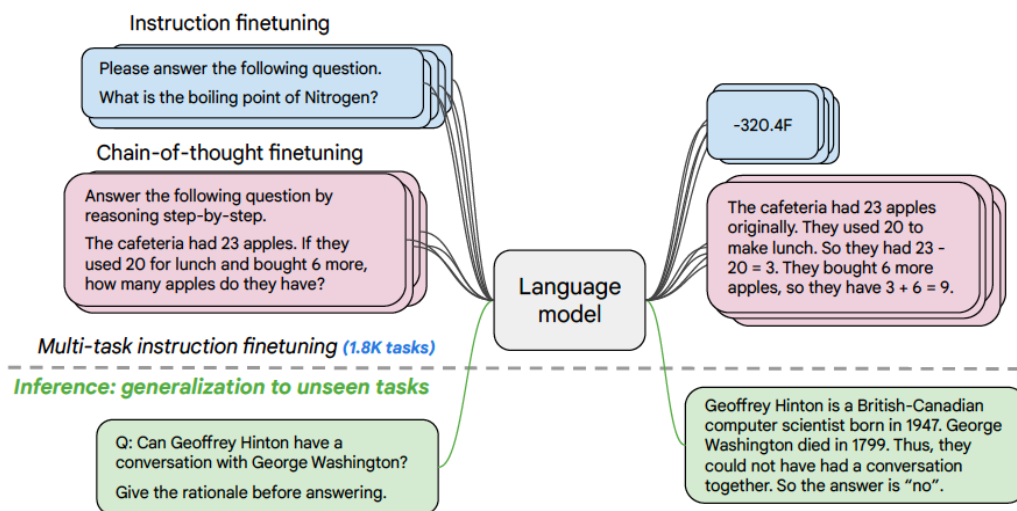
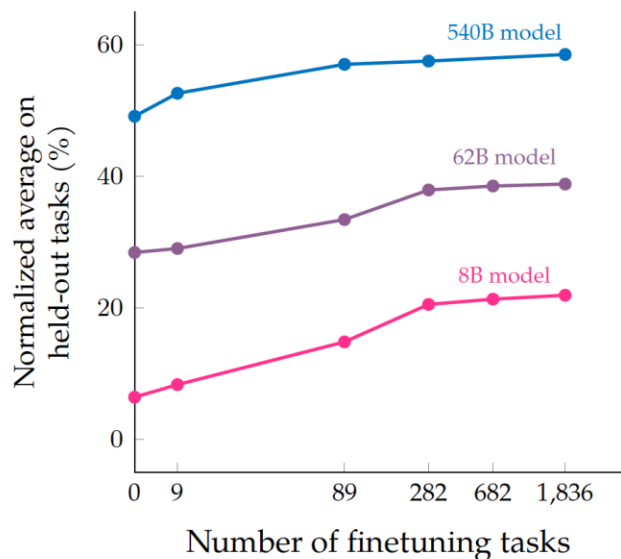


大模型真正展现了什么能力?

超强的文本建模能力： 3.2万长度Token的理解和生成能力

多任务统一学习范式： 将所有NLP任务统一到**生成式多任务学习框架**

任务泛化能力： 数千种任务统一框架学习后，部分无标注数据任务上效果很好



大模型擅长的任务

问答：开放领域问答、专业领域问答、阅读理解

开放生成：框架、续写、重写、扩展

给定内容生成：翻译、摘要生成

跨领域分类：类别分类、主题分类、倾向性分类

代码：生成代码、修改bug

信息抽取：命名实体识别、关系抽取、结构化数据提取

大模型带来了新的范式更迭

小模型时代的工作流程



提出需求



转换为
机器学习问题



产品经理



研发

标注数据



训练模型



模型评测



模型部署



N个小模型

小模型时代的难点

任务开发成本高、时间周期长

相同任务的微小需求变化，需要30%-70%的重新开发成本，**无法产品化**

模型开发和维护成本高

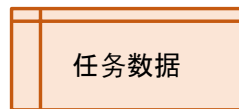
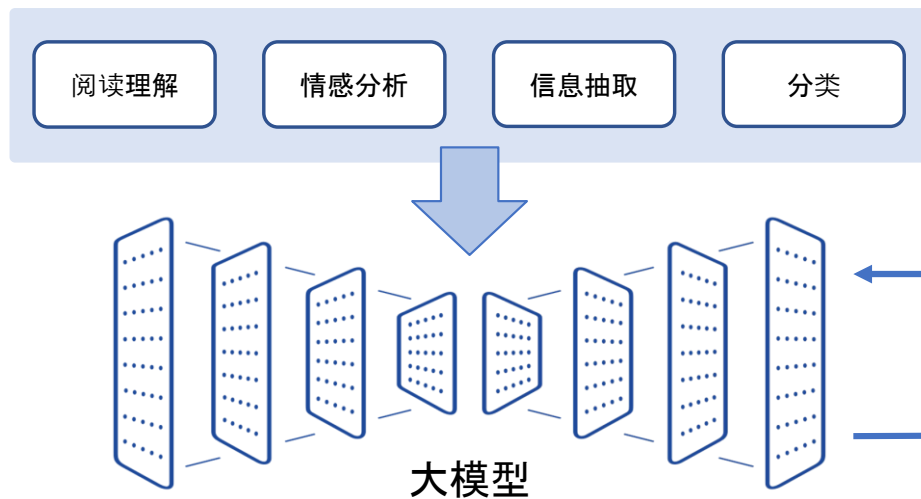
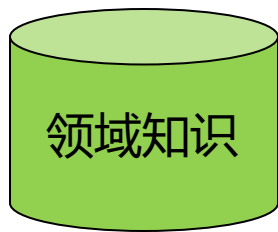
高水平算法研究员稀缺

将来会演变成什么样子?

大规模语言模型构建

通用能力注入

特定任务使用



任务指令构造

结果

任务范式发生巨大变化

使用自然语言对模型进行训练和使用

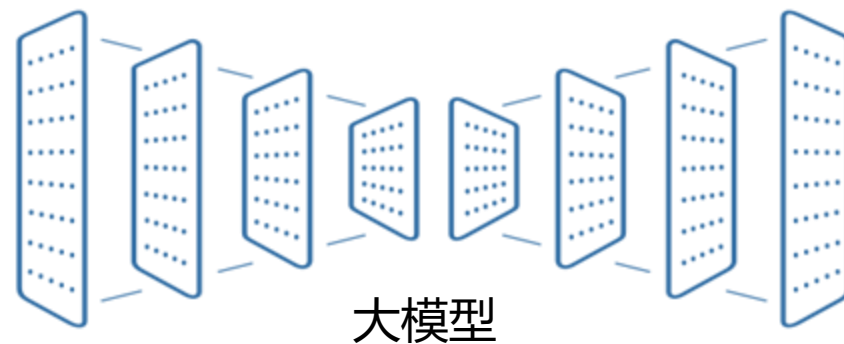
【句子输入】，上述文字表达了褒义和贬义倾向？

【文章输入】，上述文章中包含哪些公司主体？

【文章输入】，从上述文章提取2022年公司主营业务收入

【财报输入】，从上述财报，输出核心提示

【数据】，根据上述数据，书写2023年第一季度财报摘要



大模型优势

非常少量算法研究员

新任务可以**快速训练**，并且**不需要模型重新部署**

低成本产品化

可以应用于哪些场景——银行

场景一: 信贷辅助决策 业务部门: 风险/公司/信等

财务报告自动化解析
公司舆情数据智能挖掘
信贷流水材料智能处理
宏观行业景气度分析
风险事件传导分析

NLP-抽取/OCR/表格解折
NLP-分类/情感/实体识别/事件标签
NLP-表格解折/OCR
NLP-抽取/事件标签/专家知识
NLP-事件/专家知识

场景二: 公司债业务审核业务部门: 投行

募集债券说明书智能审核
审计报告自动化比对

NLP-抽取/审核/比对
NLP-抽取/比对

场景三, 智能客服业务部门: 远程银行/信用卡等

智能客户服务
工单数据处理
多元营销辅助

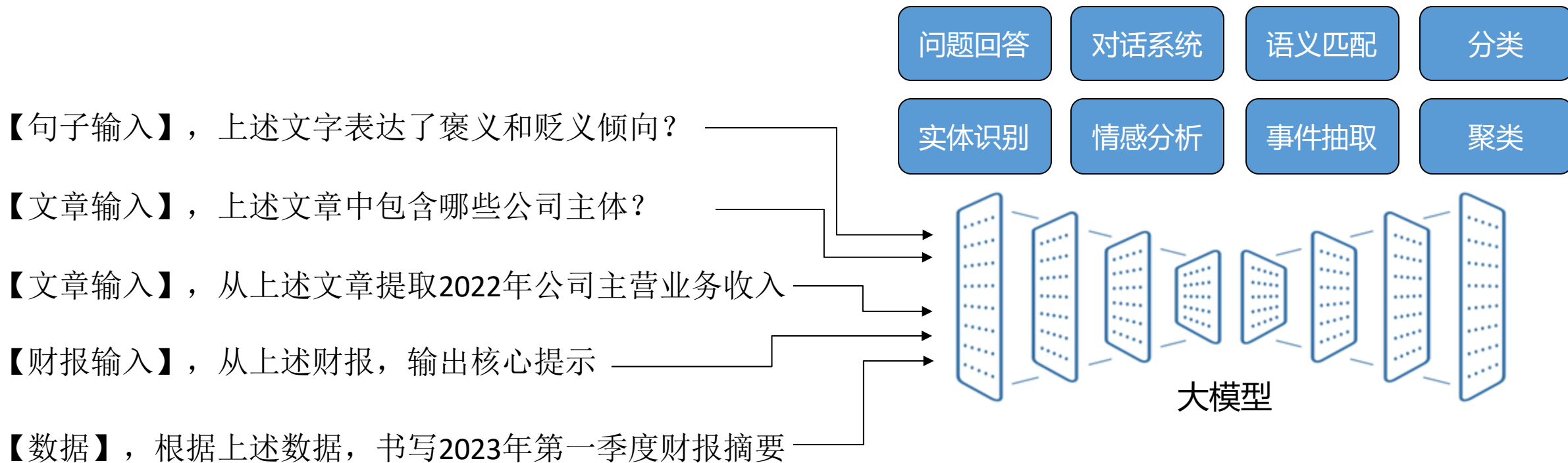
NLP-多轮对话
NLP-分类/聚类/OCR/估感
NLP-分类/聚类/实体/情感

场景四:企业知识库建设业务部门, 科技部门

企业内部数据

NLP-搜索/问答/文本解析

银行应用方式



增加一个新任务时怎么办？



甲方

提出需求



转换为
机器学习问题



产品经理

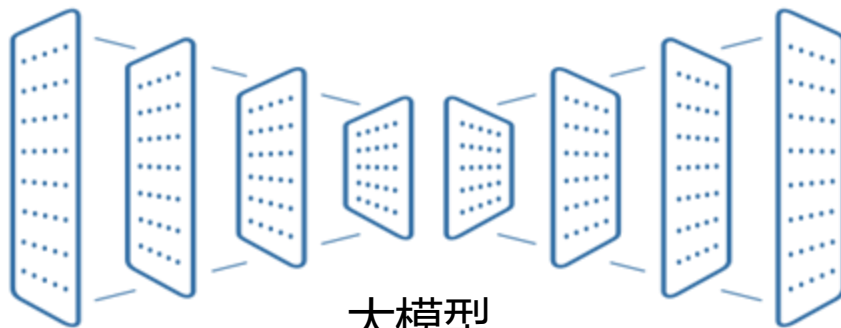
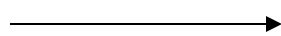


产生训练数据



研发

【句子输入】，上述文本是关于信用卡还是贷款业务？



大模型

大模型时代的自然语言处理研究变化

大模型时代从**单个小任务**的研究，转向**一类问题统一框架**研究

机遇： NLP从**手工作坊**转向**蒸汽机时代**

面向B端私有化部署的**"小"**模型

如何训练自己的领域大模型

场景

根据业务分析是否需要大模型

模型需求稳定，模型数量少，目前效果较好



模型需求稳定，但效果亟待提升



大模型的长距离建模能力，解决歧义问题

模型需求稳定、模型数量大



大模型的统一任务建模能力，解决模型模型数量多、管理难问题

小模型不能解决的问题



AIGC：对话、问题、生成等

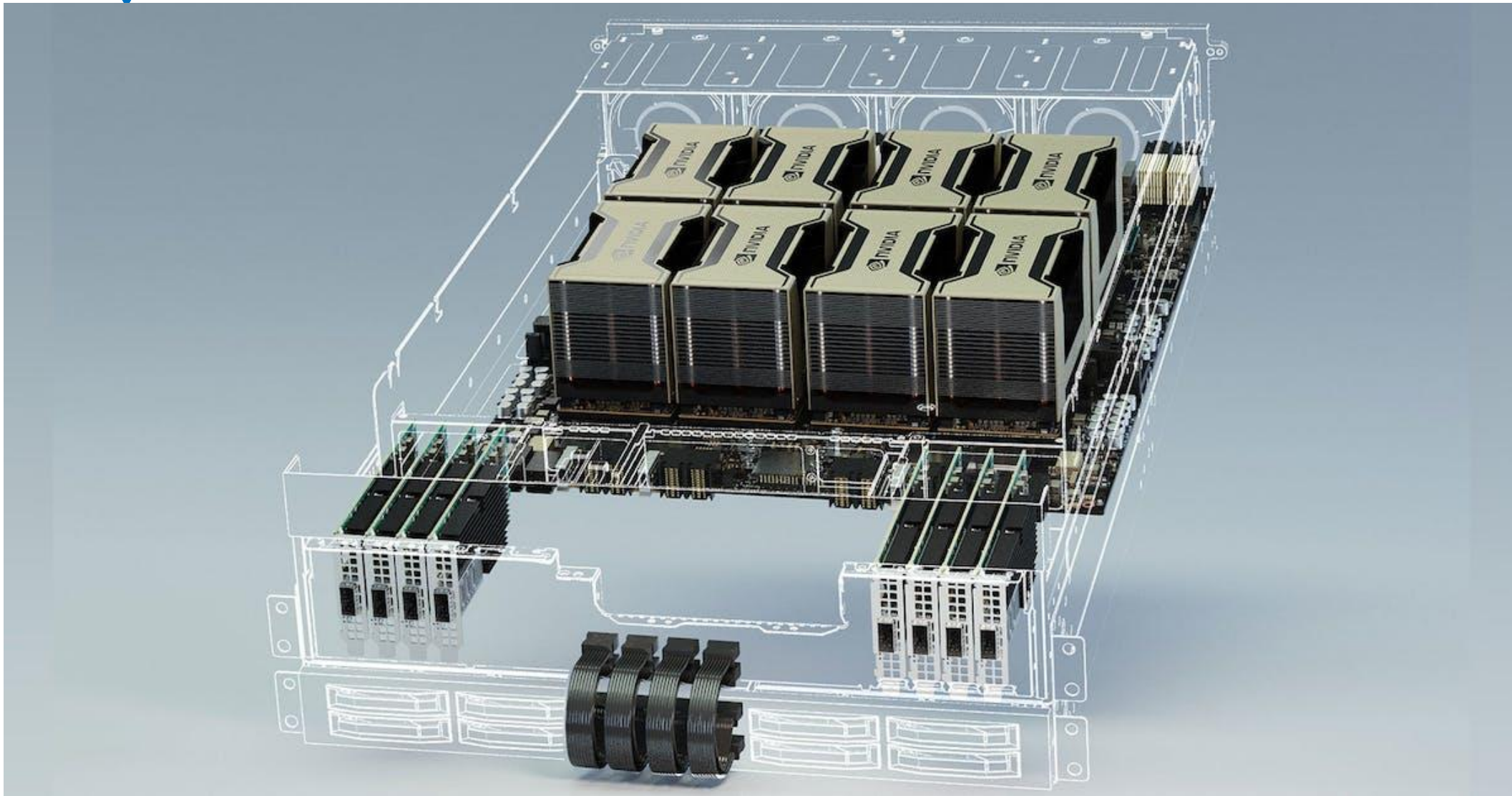
模型需求不稳定



大模型的领域和任务泛化能力，解决领域迁移问题

特别提醒：大多数情况下仍然依赖一定数量的训练数据

硬件资源



分布式训练框架



DeepSpeed

Megatron-DeepSpeed

NVIDIA/Megatron-LM



Colossal-AI

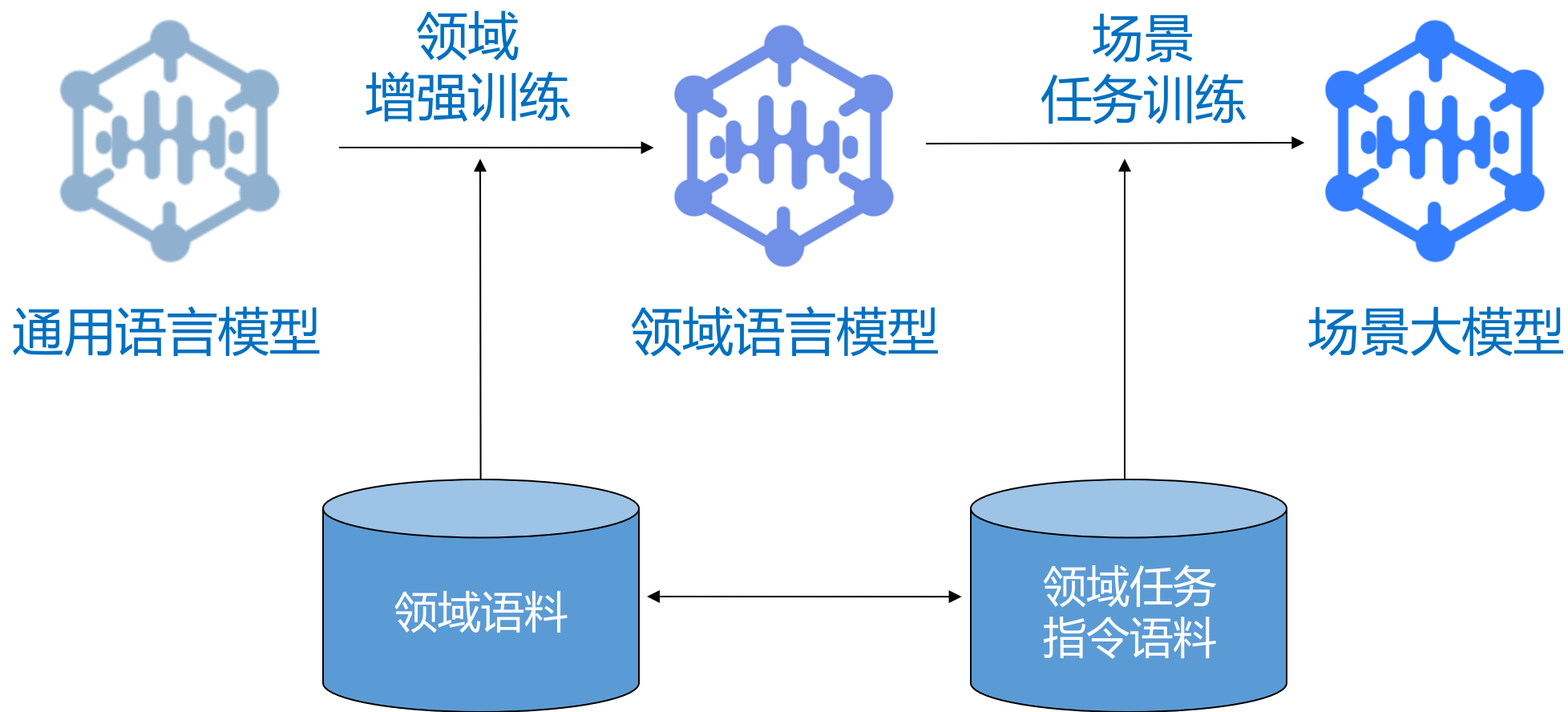


Accelerate



LMFLOW

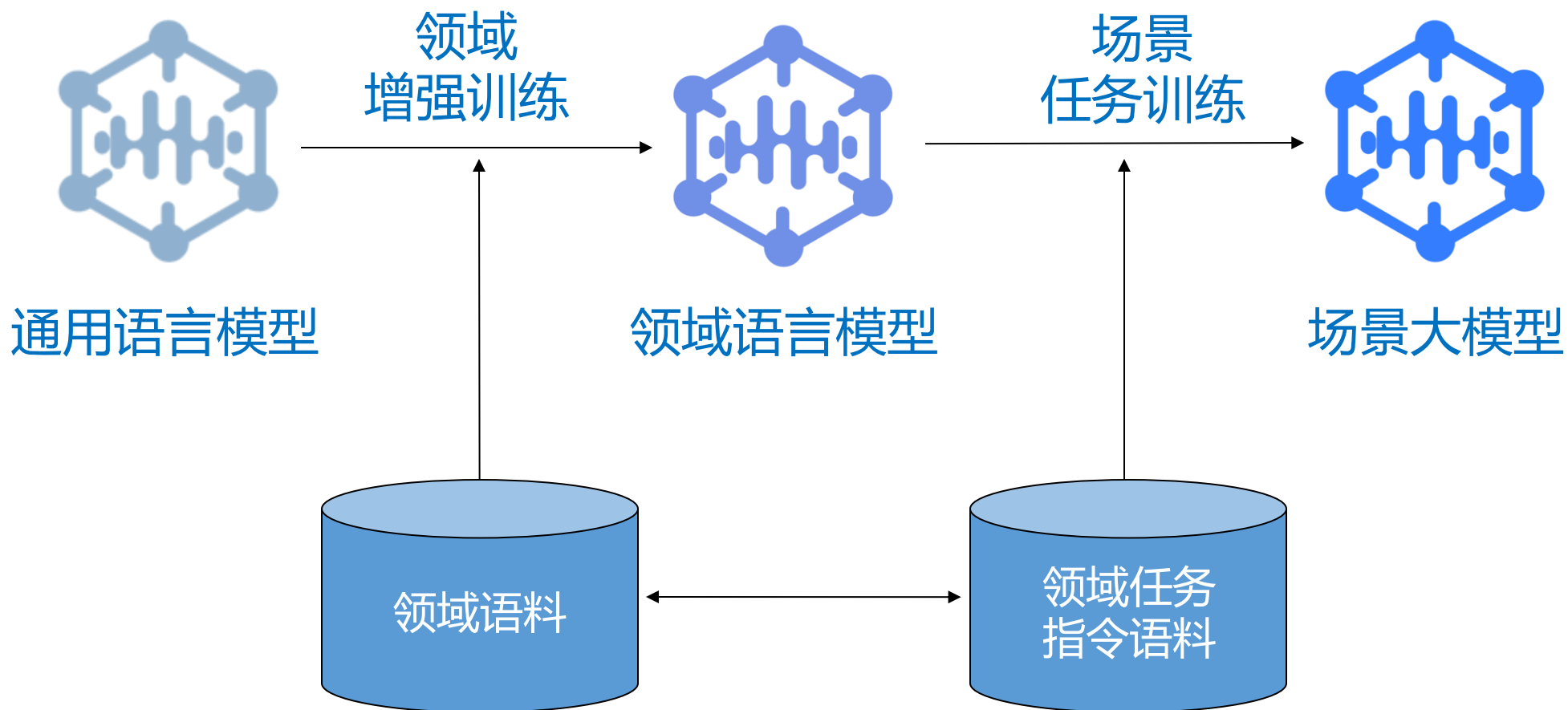
学习过程



学习过程

难点：1. 通用到场景增强训练方法
2. 场景数据收集

难点：1. 场景任务定义
2. 与现有系统的关系和对接方式
3. 训练数据构建方法
4. 场景任务训练



如何搞

大模型需要：

1. 领域语言模型，包含任务所需知识
2. 指令理解数据要满足多样性、语言模型对齐、规模等要求
3. 模型训练方法要合适

如何搞

1. 基础语言模型选择

- 中英文能力、Code能力、生成文本长度、领域知识量级

2. 任务数据构造

- 任务种类、指令构造、数据构造

3. 训练方案

- 基础模型、指令微调、强化学习

**不要神话和拟人化大模型
依然是统计机器学习范畴**

模型并不能真正地理解语言
模型只是根据某些统计规律来
连接它之前见过的片段



谢谢!