

ChatGPT

狂飙 原理剖析



ZOMI

ChatGPT Talk Overview

1. BERT 模型与 GPT 模型系列
2. 强化学习加入人类反馈 RLHF 模式
3. 强化学习 PG 和 PPO 算法
4. InstructGPT 原理深度剖析

从GPT1/2到GPT3

微调到prompt学习的过渡

GPT 系列

- 基于文本预训练的GPT-1，GPT-2，GPT-3 三代模型都是采用的以Transformer为核心结构的模型，在网络模型层面不同的是模型的层数和词向量长度等超参。

模型	发布时间	层数	Head 数	词向量长度	参数量	预训练数据量
GPT-1	2018 年 6 月	12	12	768	1.17 亿	~ 5GB
GPT-2	2019 年 2 月	48	-	1600	15 亿	~ 40GB
GPT-3	2020 年 5 月	96	96	12888	1,750 亿	~ 45TB

GPT-1 : 基于 Transformer Decoder 预训练 + 微调 Finetune

Generative Pre-Training GPT，主要指 NLP 生成式的预训练模型。训练模式分为2阶段：

1. 第 1 阶段预训练：利用语言模型 LLM 进行预训练学习
2. 第 2 阶微调：通过微调 Fine tuning 解决下游任务

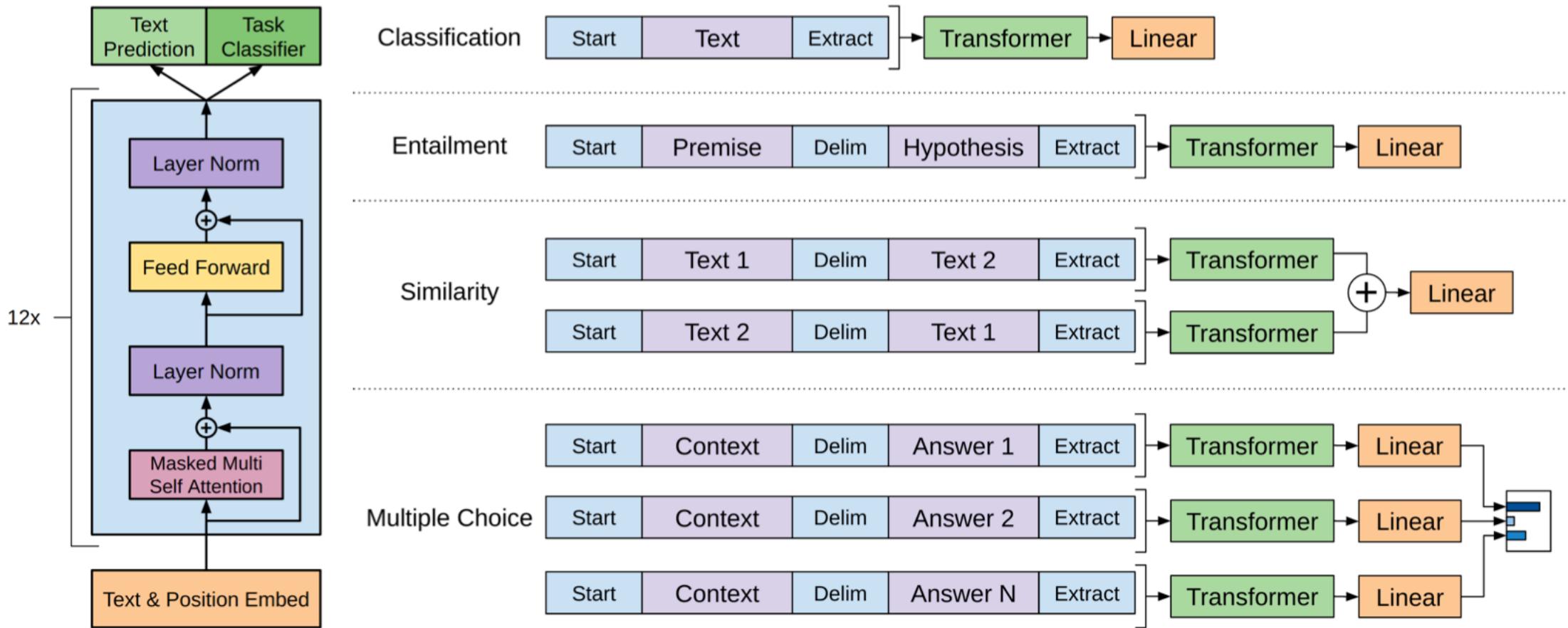
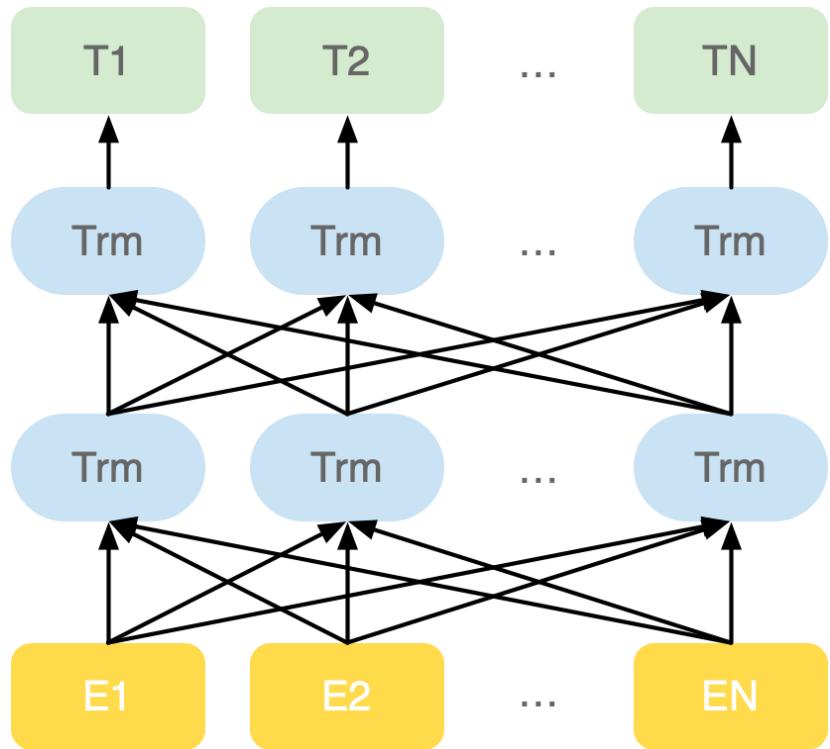


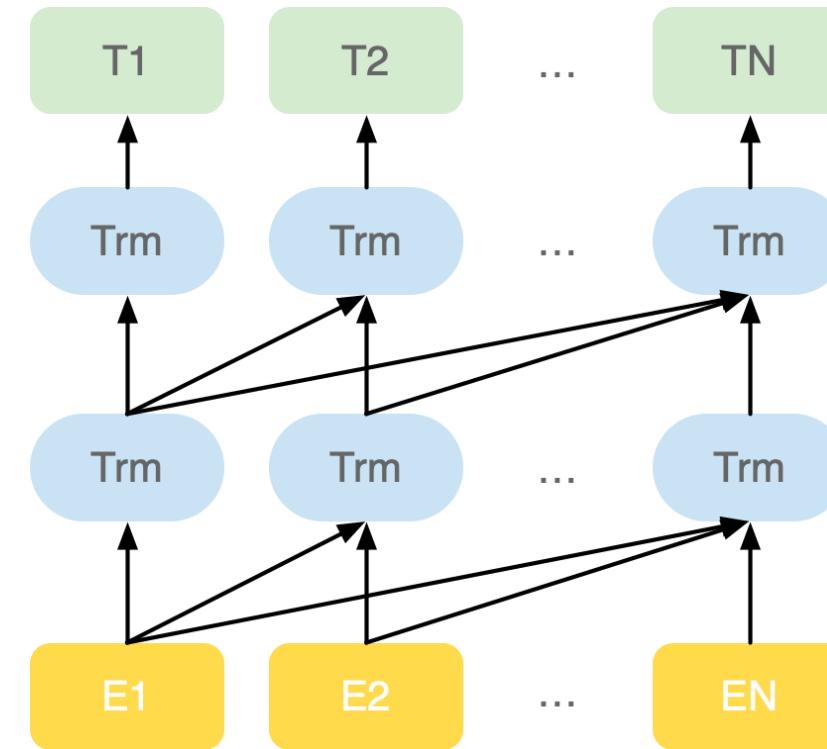
Figure 1: **(left)** Transformer architecture and training objectives used in this work. **(right)** Input transformations for fine-tuning on different tasks. We convert all structured inputs into token sequences to be processed by our pre-trained model, followed by a linear+softmax layer.

GPT 与 Bert 区别

Google BERT

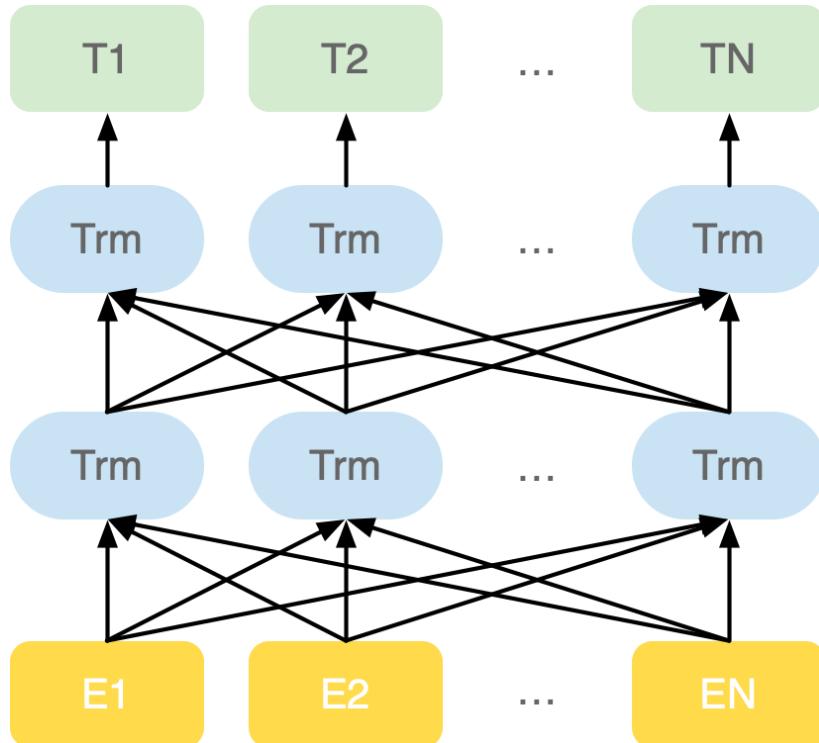


OpenAI GPT



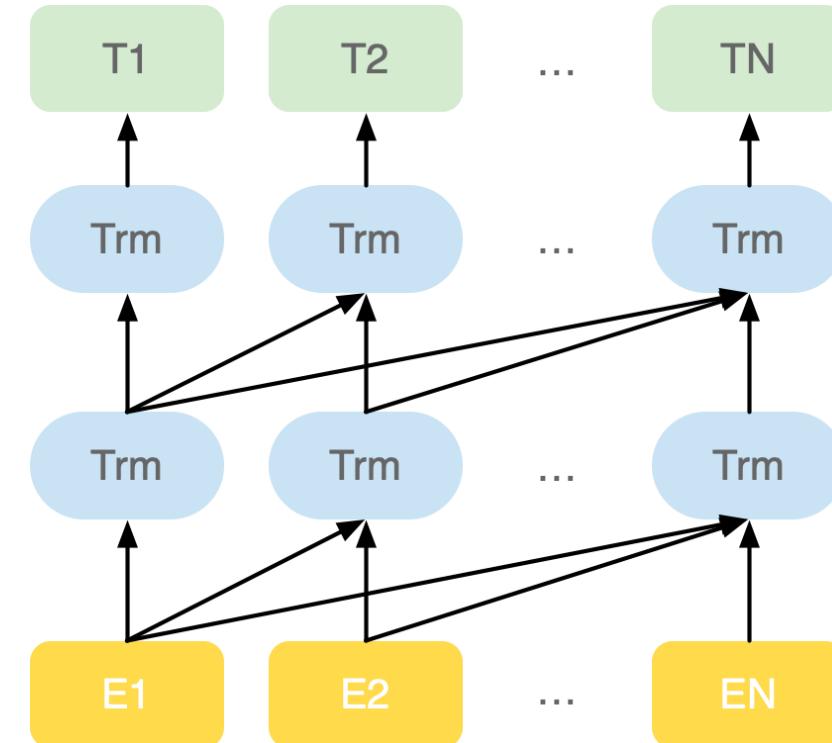
GPT 与 Bert 区别

ZOMI 经常更新 (视频) 在大晚上



ZOMI 经常更新 ____ 在大晚上

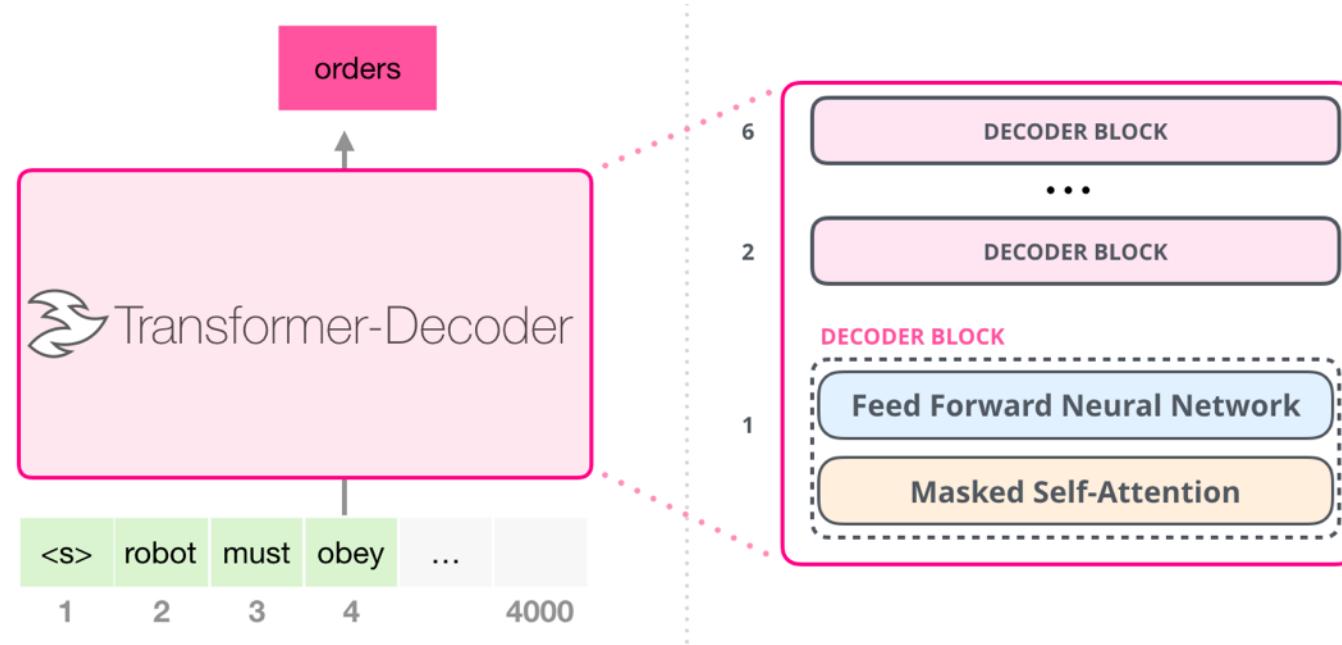
ZOMI 经常大晚上更新 (视频)



ZOMI 经常大晚上更新 ____

GPT-1：基于 Transformer Decoder 预训练 + 微调 Finetune

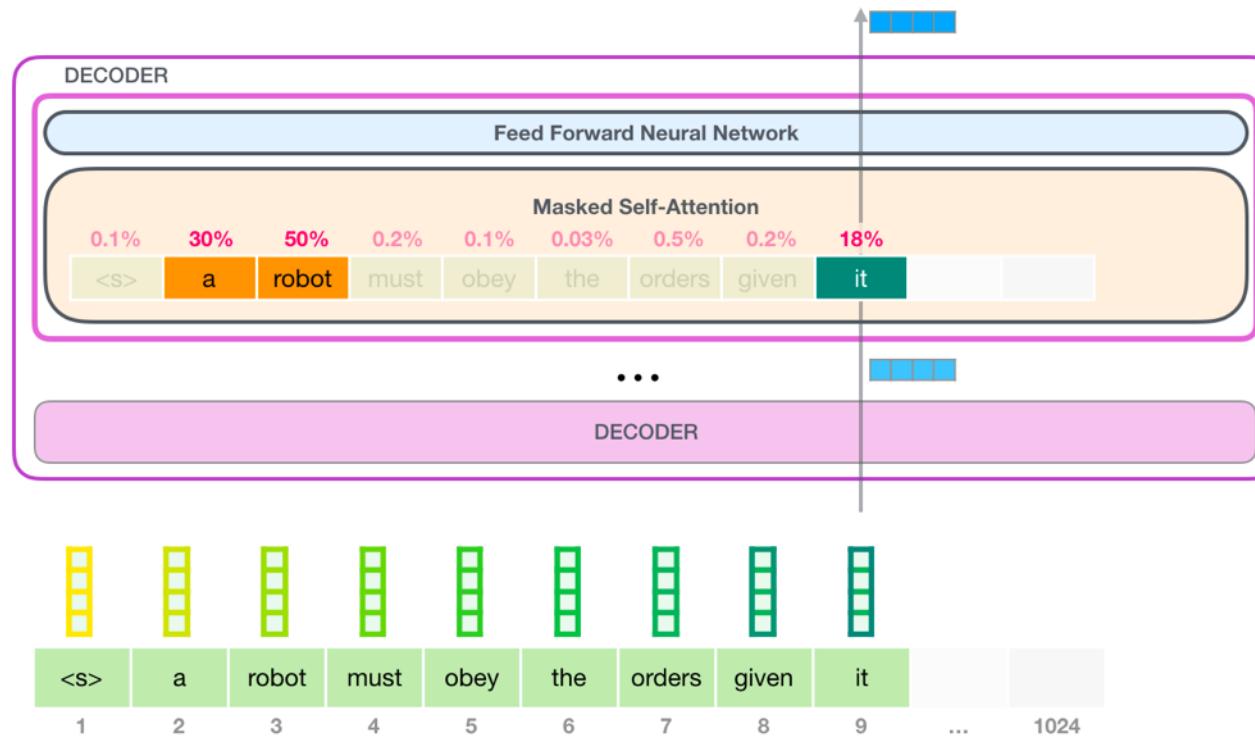
- GPT-1 预训练以语言模型作为目标任务，但是采用的是单向的语言模型
- Bert 网络结构类似于 Transformer Encoder 部分，而 GPT 类似于 Transformer Decoder 部分
- 单从模型组成部分结构上来看，其最明显的在结构上的差异为 Multi-Head-Attention 和 Masked Multi-Head-Attention



GPT-1：基于 Transformer Decoder 预训练 + 微调 Finetune

Masked Multi-Head-Attention 就是在处理当前词的时候看不到后面的词。举个可爱栗子：

- 处理 [it] 的时候，看不到 [it] 后面的词，但会关注到 [it] 前面词中的 [a, robot]，继而注意力会计算词间 [a, robot, ..., it] 的向量及其 Attention分数的加权和，即QKV的权重。



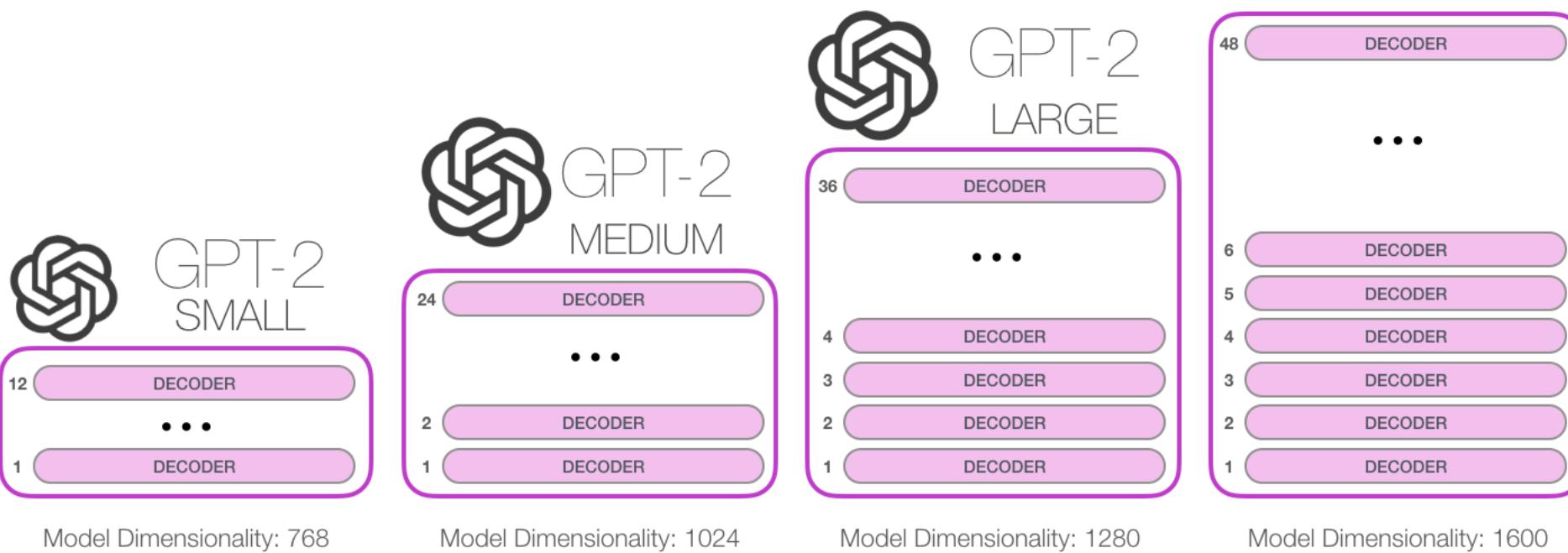
GPT-2：舍弃微调，直接利用 zero-shot learning

针对小样本/零样本的 N-shot Learning 应运而生，分为如下三种：

1. **Zero-shot Learning (零样本学习)**：没有任何训练样本进行微调训练的情况下，让预训练语言模型完成特定任务；
2. **One shot Learning (单样本学习)**：在一个训练样本进行微调训练的情况下，预训练语言模型完成特定任务；
3. **Few-shot Learning (少样本或小样本学习)**：在只有少量样本进行微调训练的情况下，预训练语言模型完成特定任务；

GPT-2：舍弃微调，直接利用 zero-short learning

- **GPT-1** 通过无标签数据上学习一个通用语言模型，再根据特定任务进行微调来处理监督任务；
- **GPT-2** 在 GPT-1 已有网络结构设计上使用了更大网络和更大数据集，并且在训练和预测过程中一次预测一个单词，以此来训练一个能够 zero-short learning 的语言模型。



GPT3：开启 NLP 新范式 prompt，实现小样本学习

Prompt-Tuning的动机旨在解决目前传统Fine-tuning的两个痛点问题：

1. **降低语义差异 (Bridge the gap between Pre-training and Fine-tuning)**：预训练任务主要以Masked Language Modeling (MLM) 为主，而下游任务则重新引入新的训练参数，因此两个阶段的目标通常有较大差异。因此需要解决如何缩小Pre-training和Fine-tuning两个阶段目标差距过大的问题；
2. **避免过拟合 (Overfitting of the head)**：由于再Fine-tuning阶段需要新引入额外的参数以适配相应的任务需要，因此在样本数量有限的情况下容易发生过拟合，降低了模型的泛化能力。因此需要面对预训练语言模型的过拟合问题。

GPT3：开启 NLP 新范式 prompt，实现小样本学习

Prompt-Tuning的动机旨在解决目前传统Fine-tuning的两个痛点问题：

1. **降低语义差异 (Bridge the gap between Pre-training and Fine-tuning)**：预训练任务主要以Masked Language Modeling (MLM) 为主，而下游任务则重新引入新的训练参数，因此两个阶段的目标通常有较大差异。因此需要解决如何缩小Pre-training和Fine-tuning两个阶段目标差距过大的问题；
2. **避免过拟合 (Overfitting of the head)**：由于再Fine-tuning阶段需要新引入额外的参数以适配相应的任务需要，因此在样本数量有限的情况下容易发生过拟合，降低了模型的泛化能力。因此需要面对预训练语言模型的过拟合问题。

GPT3 : 开启 NLP 新范式 prompt , 实现小样本学习

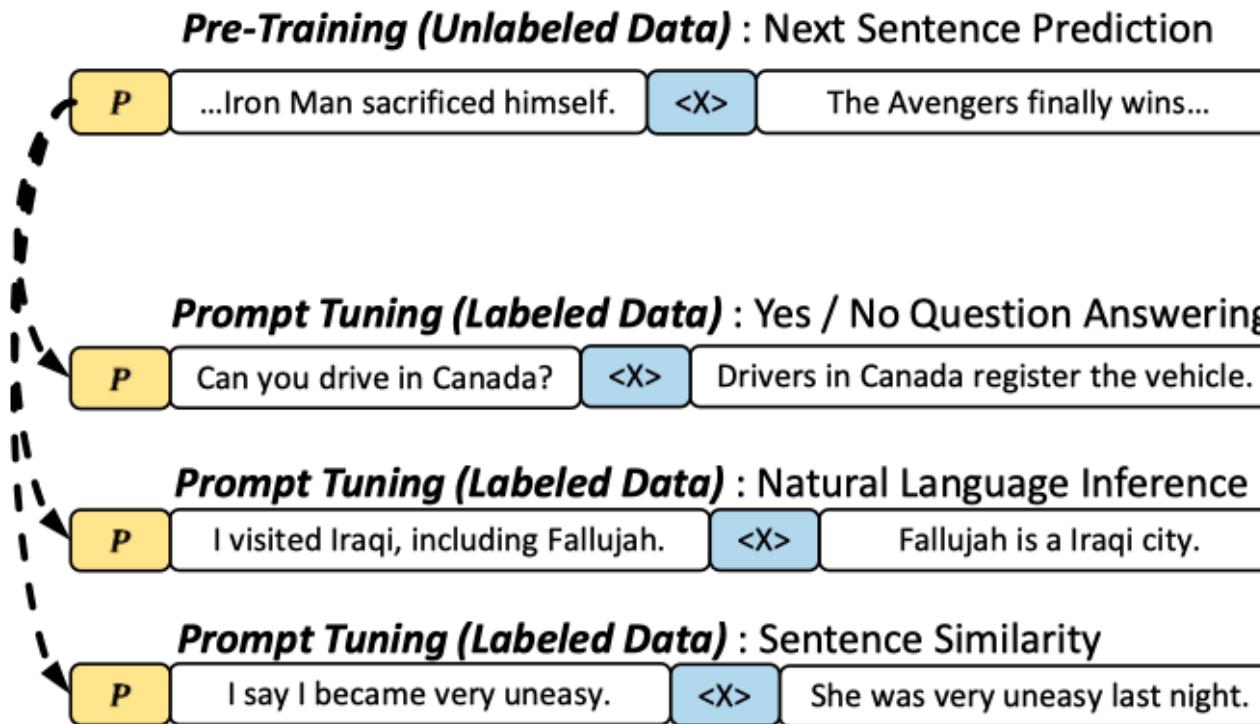


Figure 3: An example of PPT used in sentence pair tasks. P denotes soft prompt. $\langle X \rangle$ means the mask of typical encoder-decoder model like T5 and CPM-2.

GPT3 : 开启 NLP 新范式 prompt , 实现小样本学习

[example] an input that says "search" [toCode] Class App extends React Component... </div> } } }

[example] a button that says "I'm feeling lucky" [toCode] Class App extends React Component...

[example] an input that says "enter a todo" [toCode]



GPT-3



GPT 系列总结

- **Fine-tuning**：预训练语言模型通过微调适配到各种不同的下游任务。具体体现就是通过引入各种辅助任务 loss，将其添加到预训练模型中，然后继续 pre-training，以便让其更加适配下游任务。
- **Prompting**：各种下游任务能力上移到预训练语言模型。需要对不同任务进行重构，使得它达到适配预训练语言模型的效果。这个过程中，是下游任务做出了更多的牺牲，但是能够更好地统一语言模型范式。

GPT 系列总结

- 预训练 + 微调 Finetune
- 预训练 + Zero Shot
- 预训练 + Prompt Tuning
- 预训练 + Instruct Learning
- 预训练 + IL + RLHF



GPT-1

GPT-2

GPT-3

InstructGPT

ChatGPT

提示(Prompt Learning)学习 vs 指示学习(Instruction Learning)

相同点

- 指示学习和提示学习目的都是挖掘语言模型本身具备的知识。

区别点

- Prompt 是激发语言模型的**补全能力**，如根据上半句生成下半句，或是完形填空等。
- Instruct 是激发语言模型的**理解能力**，通过给出更明显的指令，让模型去做出正确的行动。

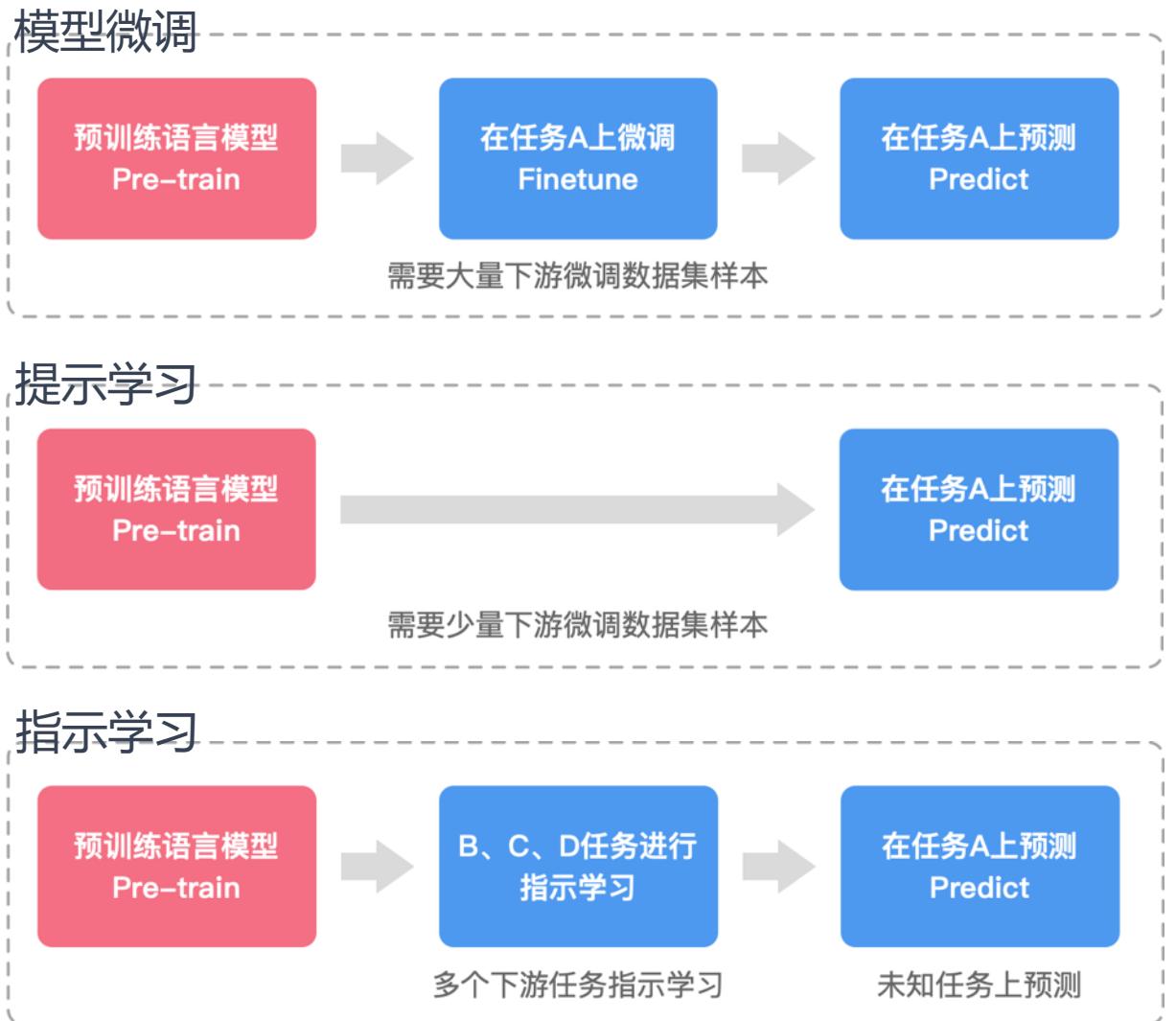
提示(Prompt Learning)学习 vs 指示学习(Instruction Learning)

举个可爱栗子：

1. 提示学习：补全这句话的内容 [给女朋友买了这个项链，她很喜欢，这个项链太_____了]
2. 指示学习：判断这句话的情感 [给女朋友买了这个项链，她很喜欢。A=好；B=一般；C=差]

提示(Prompt Learning)学习 vs 指示学习(Instruct Learning)

- 指示学习和提示学习的目的都是去挖掘语言模型本身具备的知识。不同的是Prompt是激发语言模型的**补全能力**，例如根据上半句生成下半句，或是完形填空等。Instruct是激发语言模型的理解能力，它通过给出更明显的指令，让模型去做出正确的行动。



引用

1. Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018)
2. Radford, Alec, et al. "Language models are unsupervised multitask learners." OpenAI blog 1.8 (2019): 9
3. Dale, Robert. "GPT-3: What's it good for?." Natural Language Engineering 27.1 (2021): 113-118
4. Floridi, Luciano, and Massimo Chiriatti. "GPT-3: Its nature, scope, limits, and consequences." Minds and Machines 30 (2020): 681-694.
5. Brown, Tom, et al. "Language models are few-shot learners." Advances in neural information processing systems 33 (2020): 1877-1901
6. Ouyang, Long, et al. "Training language models to follow instructions with human feedback." arXiv preprint arXiv:2203.02155 (2022)
7. <https://openai.com/blog/instruction-following/>
8. <https://openai.com/blog/chatgpt/>
9. <https://sh-tsang.medium.com/review-instructgpt-training-language-models-to-follow-instructions-with-human-feedback-7fce4bf9059a>
10. <https://jalammar.github.io/how-gpt3-works-visualizations-animations/>
11. <https://jalammar.github.io/>



BUILDING A BETTER CONNECTED WORLD

THANK YOU

Copyright©2014 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.