

订阅DeepL Pro以编辑此演示文稿。  
访问[www.DeepL.com/pro](https://www.deepl.com/pro?cta=edit-document)，了解更多信息。

**自我培训：对齐语言模式**

**带有自我生成的指令**

|  |  |  |  |
| --- | --- | --- | --- |
| **王一中** | **Yeganeh Kordi** | **Swaroop Mishra Alisa Liu** | |
| **Noah A.** Smith + | **丹尼尔-卡沙比** | | **Hannaneh** Hajishirzi + |
| 华盛顿大学德黑兰理工学院 亚利桑那州立大学 | | | |
| 约翰霍普金斯大学 | | +Allen人工智能研究所 | |

yizhongw@cs.washington.edu

|  |
| --- |
| arXiv:2212.10560v1 [cs.CL] 2022年12月20日 |

**摘要**

大型的 "指令调谐 "语言模型（为响应指令而进行微调）已经显示出对新任务进行零点生成的非凡能力。然而，它们在很大程度上依赖于人类编写的指令数据，而这些数据在数量、多样性和创造性方面都是有限的，因此阻碍了调整后的模型的通用性。我们介绍了SELF - INSTRUCT，这是一个通过引导生成预训练的语言模型的指令跟随能力的框架。我们的管道从语言模型中生成指令、输入和输出样本，然后在使用它们来微调原始模型之前对它们进行修剪。将我们的方法应用于vanilla GPT3，我们证明了在SUPER-NATURALINSTRUCTIONS上比原始模型有33%的绝对改进，与InstructGPT0011 的性能相当，后者是用私人用户数据和人类注释训练的。为了进一步评估，我们策划了一套由专家编写的新任务指令，并通过人类评估表明，用SELF-INSTRUCT调整GPT3的性能远远超过了使用现有的公共指令数据集，与InstructGPT001的绝对差距仅有5%。SELF-INSTRUCT提供了一种几乎不需要注释的方法，用于将预先训练好的语言模型与指令对齐，我们发布了我们的大型合成数据集，以促进未来对指令调整的研究2 。

* **简介**

最近的NLP文献见证了在建立模型方面的大量活动，这些模型包括

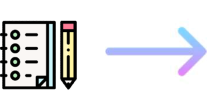
1除非另有说明，我们的比较是与text-davinci-001引擎的比较。我们专注于这个引擎，因为它最接近我们的实验设置：有监督的微调与人类演示。较新的引擎更加强大，尽管它们使用了更多的数据（如代码补全或最新的用户查询）或算法（如PPO），这些都是难以比较的。

2代码和数据将在[https://github. com/yizhongw/self-instruct](https://github.com/yizhongw/self-instruct).

可以遵循自然语言指令（[Mishra](#page12)等人，[2022](#page12)；Wei等人，[2022](#page13)；Sanh等人，[2022](#page12)；Wang等人，[2022](#page13)；Ouyang等人，[2022](#page12)；Chung等人，[2022](#page11)，i.a）。这些发展由两个关键部分驱动：大型预训练的语言模型（LM）和人类写的指令数据。PROMPTSOURCE（Bach等人，[2022](#page11)）和SUPER-NATURALINSTRUCTIONS（Wang等人，[2022](#page13)）是最近两个值得注意的数据集，它们使用大量的人工注释来收集指令来构建T0（Bach等人，[2022](#page11)；Sanh等人，[2022](#page12)）和Tk-INSTRUCT（Wang等人，[2022](#page13)）。然而，这个过程是昂贵的，而且鉴于大多数人类世代倾向于流行的NLP任务，往往suﬀ了有限的多样性，没有涵盖真正的各种任务和diﬀ不同的描述方式。鉴于这些局限性，要继续提高指令调谐模型的质量，就必须开发其他方法来监督指令调谐模型。

在这项工作中，我们引入了SELF-INSTRUCT，这是一个半自动化的过程，利用模型本身的指令信号对预训练的LM进行指令调整。整个过程是一个迭代的引导算法（见图1），它从一个有限的（例如，在我们的研究中是175条）人工编写的指令种子集开始，这些指令被用来指导整个生成过程。在第一阶段，模型被提示生成新任务的指令。这一步是利用现有的指令集来创建更广泛的指令，以定义（通常是新的）任务。考虑到新生成的指令集，该框架还为它们创建了输入输出实例，这些实例随后可用于监督指令的调整。最后，在将其添加到任务池之前，会使用各种措施来修剪低质量和重复的指令。这个过程可以重复许多次互动，直到达到大量的任务。

为了实证评估SELF-INSTRUCT，我们运行了



**175个种子任务有**

**1个指令和**

**每个任务有1个实例**

**任务库**



**LM**



**任务**

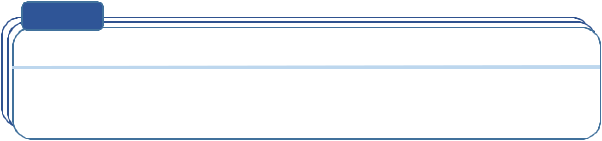


**第1步：指令生成**



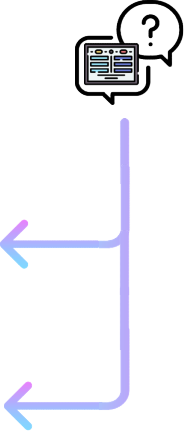
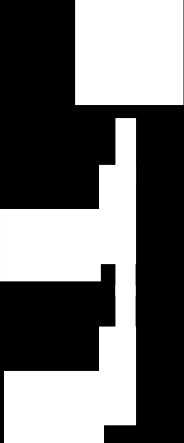
**任务**

|  |  |
| --- | --- |
| **指示：给我一个报价，从一个** | **LM** |
|  |
| **在这个问题上的著名人士。** |  |
| **第3步：实例生成** |  |



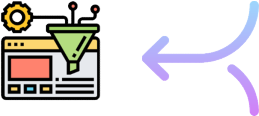
**第2步：分类**

**任务识别**



**是**

**第4步：过滤**



**指示：找出所给文本是支持还是反对堕胎。**

**类标签：支持堕胎**

**输入：文本：我认为妇女应该有权选择是否执行任务**

**他们是否想进行堕胎。**



**任务**

**指示：给我一个关于这个话题的名人名言。**

**输入：主题：诚实的重要性。**

**任务**

**输出："诚实是智慧之书的第一章"。- 托马斯-杰斐逊**

**输出优先** **LM**



**没有**

**输入优先**

图1：SELF-INSTRUCT的高层概述。这个过程从一个小型的种子任务集（每个任务有一个指令和一个输入输出实例）作为任务池开始。随机任务从任务池中被抽出，用来促使一个现成的LM生成新的指令和相应的实例，然后过滤低质量或类似的世代，再加回初始任务库。由此产生的数据可以用于语言模型本身的指令调整，以便更好地遵循指令。图中所示的任务是由GPT3生成的。更多创造性的例子[见表10](#page18)。

|  |  |  |
| --- | --- | --- |
| 这个框架对GPT3（Brown等人，2020）、 | (2) 我们通过扩展的数据来证明其有效性。 | |
| 这是一个虚拟的LM（[§](#page5)4）。迭代的SELF- | (3)我们发布了一个关于 "在中国 "的信息。 | |
| 在这个模型上的INSTRUCT过程导致了大约 | 一个包含52K指令的大型合成数据集和一个 | |
| 52K指令，与约82K实例配对 | 一组手动编写的新任务，用于构建和 | |
| 输入和目标输出。我们观察到，结果- | 评估未来的指令跟随模式。 | |
| 摄取数据提供了一个多样化的创造性任务 | **2 相关工作** |  |
| 而其中50%以上的人的ROUGE-小于0.3。 |  |
| L与种子指令重合（§4[.2](#page5)）。关于 | 许多作品都发现了香草语言的证据。 | |
| 这些数据，我们建立了GPT3SELF[-I](#page5)NST 构建 |
| 对GPT3进行微调（即，用于GPT3的同一模型 | **指示跟随的语言模型。** | A- |
| 仪表模型在以下情况下是有效的 | |
| 生成教学数据）。 我们评估 |
| 语言说明，如果用注释的 "在-"来调整。 | |
| GPT3 SELF-INST ，与其他各种模式相比。 |
| 结构性 "数据--包含语言的数据集 | |
| 他在SUPER-Learning中包含的典型NLP任务上的表现都很好。 | 教学命令及其预期结果 | |
| 自然指南》（[Wang等人](#page13)，[2022年](#page13)），以及《自然指南》（NaturalINSTRUCTIONS）。 |
| 基于人类的判断（[Weller等人](#page13)，[2020](#page13)； | |
| 一套新的指令，是为新奇的事物而创建的。 |
| [Mishra等人](#page12), [2022](#page12); [Wang等人](#page13), [2022](#page13); [Wei等人](#page13)、 | |
| 使用指令跟随模型（§[5](#page5)）。的 |
| [2022](#page13); [Sanh等人](#page12), [2022](#page12); [Ouyang等人](#page12), [2022](#page12); Par[-](#page12) | |
| SUPERNI的结果表明，GPT3 SELF- I NST 出。 |
| [mar等人](#page12)，[2022](#page12)；[Scialom等人](#page12)，[2022](#page12)；[Chung等人](#page11)、 | |
| 在很大程度上胜过了GPT3（原始模型）。 | [2022](#page11); [Luo et al](#page12)., [2022](#page12); [Puri et al.](#page12), [2022](#page12); [Yin et al](#page13)、 | |
| 的利润率(+33.1%)，并几乎达到了 |
| [2022](#page13)；[Chakrabarty等人](#page11)，[2022](#page11)；[Lin等人](#page12)，[2022](#page12)； | |
| InstructGPT的表现001 。此外，我们的人类 |
| [Gupta等人](#page11)，[2022](#page11)；[Muennighoﬀ等人](#page12)，[2022](#page12)）。加 -- | |
| 对新创建的指令集进行评估 | 从理论上讲，它们显示出了一个直接的相关性。 | |
| 显示，GPT3SELF-INST 展示了广泛的 |
| 规模和多样性的 "教学 "数据和 | |
| 能力后的教学范围，优于 | 形成的模型对未见过的任务的可推广性。 | |
| 在其他可公开获得的数据基础上训练的模型。 |
| 由于这些发展取决于人的年报。 | |
| 构建数据集，只留下5%的差距 |
| 这就造成了一个瓶颈。 | |
| InstructGPT001 。 |
| 为实现更多的通用模型的进展（为 | |
| 综上所述，我们的贡献是：(1) 自我 | 例子见[Wang等人](#page13)，[2022年](#page13)的图5a）。我们的 | |
| INSTRUCT，一种诱导教学的方法--。 | 这项工作的目的是解决这一瓶颈问题，通过减少 | |
| 用最少的人类标记来跟踪能力 | 对人类注释者的依赖性。 |  |

此外，尽管像InstructGPT（Ouyang等人，[2022年](#page12)）这样的模型表现出色，但其构建过程仍然相当不透明。特别是，由于这些关键模型背后的主要企业实体发布的透明度和数据有限，*数据*的作用仍然没有得到充分研究。要解决这样的挑战，就必须建立一个涵盖广泛任务的大规模公共数据集。

在多模态学习文献中，指令跟随模型也备受关注（[Fried](#page11)等人，[2018](#page11)；Shridhar等人，[2020](#page13)；Min等人，[2022](#page12)；Weir等人，[2022](#page13)）。SELF-INSTRUCT作为扩展数据的一般方法，有可能在这些环境中也有帮助；然而，这不在这项工作的范围之内。

**用于数据生成和增强的语言模型。**许多工作都依靠生成型LM来生成数据（Schick和[Schütze](#page12)，[2021](#page12)；Wang等人，[2021](#page13)；Liu等人，[2022](#page12)；Meng等人，[2022](#page12)）或增强（Feng等人，[2021](#page11)；Yang等人，[2020](#page13)；Mekala等人，[2022](#page12)）。例如，Schick和Schütze（[2021年](#page12)）提议用提示性的大型LM取代人类*对某项任务的*注释，并在SuperGLUE任务的背景下使用所产生的数据来微调（通常较小的）模型（Wang等人，[2019年](#page13)）。虽然我们的工作可以被视为一种 "增强 "的形式，但我们的工作与这一路线不同，因为它*不是*针对某个特定的任务（例如，QA或NLI）。相比之下，SELF-INSTRUCT的一个明显的动机是引导新的任务定义，这些任务可能以前没有被任何NLP从业者定义过（尽管可能对下游用户仍然很重要）。

**自我训练。**一个典型的自我训练框架（He等人，[2019](#page11)；Xie等人，[2020](#page13)；Du等人，[2021](#page11)；Amini等人，[2022](#page11)；Huang等人，[2022](#page11)）使用训练好的模型来给未标记的数据分配标签，然后利用新标记的数据来改进该模型。在类似的路线中，Zhou等人（[2022a](#page13)）使用多个提示来指定一个任务，并建议通过提示的一致性来规范化，鼓励对提示进行一致的预测。这允许用额外的未标记的训练数据对模型进行微调，或者在推断时直接应用。虽然SELF- INSTRUCT与自我训练文献有一些相似之处，但大多数自我训练方法都假定有一个特定的*目标任务*以及该任务下的*未标记的例子*；相比之下、

自主研发，从零开始生产各种任务。

**知识提炼。**知识蒸馏（Hinton等人，[2015](#page11)；Sanh等人，[2019](#page12)；[West](#page13)等人，[2021](#page13)；Magister等人，[2022](#page12)）经常涉及到知识从大模型转移到小模型。SELF-INSTRUCT也可以被看作是 "知识蒸馏 "的一种形式，然而，它在以下方面与这一路线不同：（1）蒸馏的来源和目标是相同的，即一个模型的知识被蒸馏到自己身上；（2）蒸馏的内容是指令任务的形式（即定义一个任务的指令，和一组实例）。

**用有限的资源进行引导。**最近的一系列工作使用语言模型，用专门的方法引导一些推断。NPPrompt（Zhao等人，[2022](#page13)）提供了一种无需任何微调就能生成语义标签预测的方法。它使用模型自身的嵌入来自动寻找与数据样本的标签相关的词，因此减少了对模型预测到la-bel（言语者）的人工映射的依赖。STAR（Zelikman等人，[2022年](#page13)）反复利用少量的推理例子和没有推理的大数据集，来引导模型进行推理的能力。自我修正（Welleck等人，[2022](#page13)）将一个不完美的基础生成器（模型）与一个单独的修正器解耦，该修正器学会了迭代修正不完美的生成器，并展示了对基础生成器的改进。我们的工作则侧重于在指令范式中引导新任务。

**指令生成。**最近的一系列工作（Zhou等人，[2022b](#page13)；Ye等人，[2022](#page13)；[Singh](#page13)等人，[2022](#page13)；Honovich等人，[2022](#page11)）根据一些例子生成任务的指令。虽然SELF-INSTRUCT也涉及指令的生成，但我们的主要区别在于它是任务无关的；我们从头开始生成新的任务（指令和状态）。

**3** **方法**

注释大规模的指令数据对人类来说是个挑战，因为它需要：1）创造性地提出新的任务；2）为每个任务编写标记的实例的专业知识。在本节中，我们将详细介绍我们的SELF-INSTRUCT的过程，它指的是生成具有以下特征的任务的管道

一个**虚构的预训练语言模型**本身和

然后用这些生成的数据进行指令调整，以便使语言模型更好地遵循指令。这个管道在图1中被描述出来。

**3.1** **定义指令数据**

我们要生成的指令数据包含一组指令^*It*'，每条指令都用自然语言定义了一个任务*t*。每个任务有一个或多个输入-输出实例.*X ;* Yt*/*。鉴于任务指令*I*和实例输入*x，*一个模型*M*被期望产生*输出y*：*M*.*I ; x/* = *y;* for .*x; y*/*t* ¸ .*Xt;* Yt*/*。注意，在许多情况下，*指令*和实例输入并没有严格的边界。例如，"写一篇关于学校安全的文章 "可以是一个有效的指令，我们希望模型能够直接响应，而它也可以被表述为 "写一篇关于以下主题的文章 "作为指令，而 "学校安全 "作为一个实例输入。为了鼓励数据格式的多样性，我们允许这种不需要额外输入的指令（即*x*为空）。

**3.2** **自动生成指令数据**

我们生成指令数据的管道由四个步骤组成：1）指令生成，2）识别指令是否代表分类任务，3）用输入优先或输出优先的方法生成实例，4）过滤低质量数据。

**指令生成。**SELF-INSTRUCT是基于一个发现，即大型预训练的语言模型在遇到上下文中的一些现有指令时，可以被提示生成新的和新颖的指令。这为我们提供了一种从一小部分人类编写的指令种子中增长指令数据的方法。我们建议以引导的方式产生一个多样化的指令集。我们用我们的作者编写的175个任务（每个任务有1条指令和1个实例）启动了任务池。对于每一步，我们从这个池子里抽出8条任务指令作为内涵的例子。在这8条指令中，6条来自人类编写的任务，2条来自前面步骤中的模型生成的任务，以促进多样性。提示模板显示在表6中。

**分类任务识别。**因为我们需要两种不同的方法来处理分类和非分类任务，所以我们接下来要识别生成的指令是否代表分类

任务或不是。3我们用种子任务中的12条分类指令和19条非分类指令来提示香草GPT3的几率，以确定这一点。提示模板显示在表7中。**实例生成。**考虑到指令及其任务类型，我们为每个内建任务独立生成实例。这很有挑战性，因为它要求模型理解目标任务是什么，基于指令，找出需要哪些额外的输入字段并生成它们，最后通过产生输出来完成任务。我们发现，当用其他任务中的指令-输入-输出的上下文例子提示时，预训练的语言模型可以在很大程度上实现这一点。一个自然的方法是**输入优先法**，我们可以要求语言模型先根据指令想出输入字段，然后再产生相应的输出。这种生成顺序类似于模型用来响应指令和输入的方式，但这里有来自其他任务的内涵例子。提示模板显示在表8中。

然而，我们发现这种方法可以生成偏向一个标签的输入，特别是对于分类任务（例如，对于语法错误检测，它通常会生成语法输入）。因此，我们另外提出了一个用于分类任务的 "**输出优先 "方法**，即我们首先生成可能的类标签，然后根据每个类标签确定输入生成的条件。提示模板显示在[表9](#page17)中。4我们对前一步确定的分类任务采用输出优先的方法，对其余非分类任务采用输入优先的方法。

**筛选和后处理。**为了鼓励多样性，只有当一条新指令与任何现有指令的ROUGE-L重叠度小于0.7时，才会被添加到任务池中。我们还排除了包含一些特定关键词的指令（例如，图像、图片、图表），这些关键词通常不能被语言模型处理。在为每条指令生成新的实例时，我们过滤掉那些完全相同的实例或那些具有相同输入但输出不同的实例。

更具体地说，我们把具有有限的、小的输出标签空间的任务视为分类任务。

4在这项工作中，我们使用一组固定的种子任务来提示实例的生成，因此在一轮中每个任务只生成少量的实例。未来的工作可以使用随机抽样的任务来促使模型在多轮中生成更多的实例。

**3.3** **对LM进行微调以遵循指令**

在创建了大规模的指令数据后，我们用这些数据来微调原始语言模型（即SELF-INSTRUCT）。为了做到这一点，我们将指令和实例输入作为一个提示，并训练模型以标准的监督方式生成实例输出。为了使模型对不同的格式具有鲁棒性，我们使用多个模板来对指令和实例输入进行编码。例如，指令的前缀可以是 "任务："，也可以不是，输入的前缀可以是 "输入："，也可以不是，"输出："可以附加在提示的末尾，中间可以有不同数量的中断线，等等。

* **来自GPT3的自我介绍数据**

在本节中，我们将我们诱导指令数据的方法应用于GPT3作为案例研究。我们使用最大的GPT3语言模型（"davinci "引擎），通过OpenAI API访问5 。进行查询的参数在附录A.1中描述。这里我们介绍一下生成的数据的概况。

**4.1** **统计数据**

表1描述了生成的数据的基本统计数据。我们总共生成了超过52K个指令，以及超过82K个与这些指令相对应的过滤后的实例。

|  |  |
| --- | --- |
|  |  |
| 统计资料 |  |
| # 指示的数量 | 52,445 |
| - #分类指示 | 11,584 |
| - ＃非分类指示 | 40,861 |
| # 实例的数量 | 82,439 |
| - 具有空输入的实例的#个 | 35,878 |
| 平均指令长度（以字为单位） | 15.9 |
| 非空的输入长度（字数）。 | 12.7 |
| 平均输出长度（以字为单位） | 18.9 |

表1：通过对GPT3应用SELF- INSTRUCT生成的数据的统计。

**4.2** **多样性**

为了研究产生了哪些类型的指令以及这些指令的多样性，我们确定了产生的指令中的动词-名词结构。我们使用伯克利神经解析器6 （[Kitaev和Klein](#page12)，[2018](#page12)；Kitaev等人，[2019](#page11)）来解析指令，然后提取最接近于根的动词

5<https://openai.com/api/>

6<https://parser.kitaev.io/>

解析树，以及它的第一个直接名词对象。在52,445条指令中，有26,559条包含这样的结构；其他指令通常包含更复杂的子句（例如，"对这条推文是否包含政治内容进行分类。"）或被框定为问题（例如，"这些语句中哪些是真的？"）。我们在图2中绘制了前20个最常见的根动词和它们的前4个直接名词对象，这占了整个集合的14%。总的来说，我们在这些指令中看到了相当多样化的意图和文本格式。

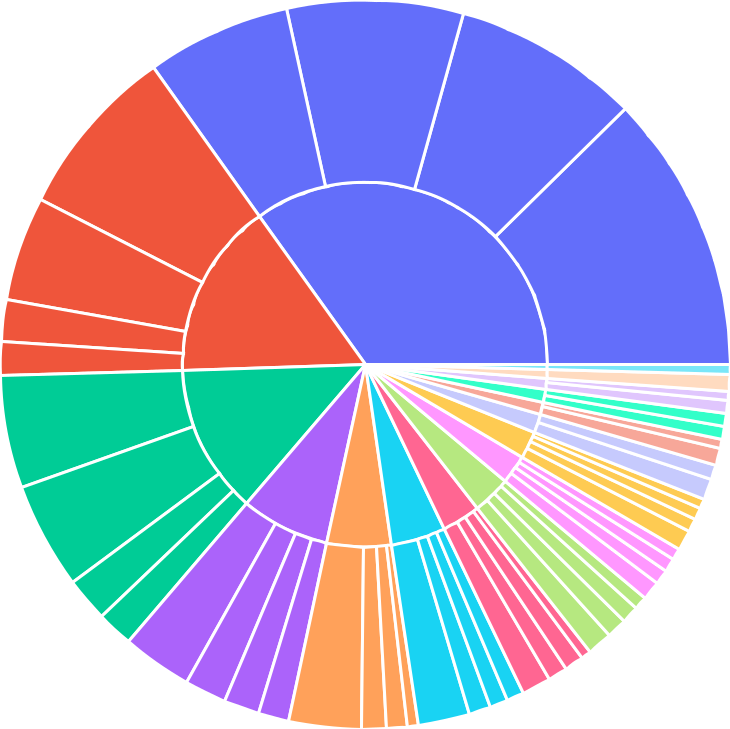
我们进一步研究生成的指令与用于提示生成的种子指令有何不同。对于每个生成的指令，我们计算其与175条种子指令的最高ROUGE-L重叠度。我们在图3中绘制了这些ROUGE-L分数的分布情况，表明有相当数量的新指令与种子指令没有太多的重叠。我们还在图4中展示了指令、实例输入和实例输出的长度多样性。

**4.3** **质量**

到目前为止，我们已经展示了生成数据的数量和多样性，但其质量仍然是不确定的。为了研究这个问题，我们随机抽取了200条指令，并在每个结构中随机选择一个实例。我们请一位专家注释者（这项工作的共同作者）从指令、实例输入和实例输出的角度来标注每个实例是否正确。表2的评估结果显示，大多数生成的指令是有意义的，而生成的实例可能包含更多的噪音（在合理的范围内）。然而，我们发现，即使生成的指令可能包含错误，但大多数指令的格式仍然是正确的，甚至是部分正确的，这可以为训练模型遵循指令提供有用的指导。我们在表10和表11中分别列出了一些好的世代和坏的世代。

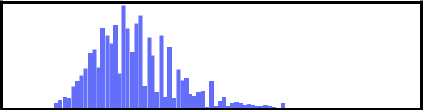
**5** **实验结果**

我们进行实验来衡量和比较各种指令调整设置下的模型质量。我们首先描述了我们的模型和其他基线，然后是我们的实验。



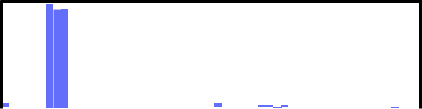
|  |  |  |  |
| --- | --- | --- | --- |
|  | 段 | 信 | 作文 |
|  |  |
|  | 例子 |  |  |
| 列表 |  | 写 | 功能 |
|  | 给 |  |
| 设置 |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 说明 | 3000 |  |  |  |  |  |
| 1000 |  |  |  |  |  |
|  | 2000 |  |  |  |  |  |
| # | 0 |  |  |  |  |  |
|  | 0.2 | 0.4 | 0.6 | 0.8 | 1 |
|  | 0 |



ROUGE-L与最相似的种子指令重合

图3：生成的指令与它们最相似的种子指令之间的ROUGE-L得分分布。



6000

建议

词

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 数 | | | | |  |
|  |  |  |  | e | |
|  |  |  | c | |  |
|  |  | n | |  |  |
|  |  | e |  |  | 节目 |
|  | t | |  |  |
| n | |  |  |  |  |
| e |  |  |  |  |  |
| s | 途径 | | |  |  |
|  |  |  |
|  |  |  |  |  | 列表 |

发现

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | 创建 |
|  |  | hm | 关于 |
|  | 它 | |
| r | |  |  |
| 去 |  |  | 小费 |
| ↪񍛨 |  |  | c |
|  |  | u |
|  |  |  | n |
|  |  |  | f |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  | 探测 | |  |  |  |  |  |  |  |  |  |  |  | 冷嘲热讽 |
|  |  |  |  |  |  |  |  |  |  |  |  |  | 预测 | |  |  |  |  |  |  |  |  |  | 感情 | | |
|  |  |  |  |  |  |  |  |  |  |  |  |  | 产量 | |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | 意识形态 | |  |  |  |  |  |  |  |  |  |  |  |  | 词 |
|  |  |  |  |  |  |  |  |  |  |  |  |  | 检验 | |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  | 告知 | |  |  |  |  |  |  |  |  |  |  | 数 | |
|  |  |  |  |  |  |  |  |  |  | 解释 | | | |  |  |  |  |  |  |  |  |  |  |  |  | 主题 |
|  |  |  |  |  |  |  |  |  | 哈 | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 哨兵 | | |
|  |  |  |  |  |  | c |  |  |  |  |  | v | |  |  |  |  |  |  |  |  |  |  |  |  | 诱骗 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 笑话 |
|  |  |  | ge | | | l | |  |  |  |  |  | e |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | a | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 故事 | |
|  |  |  |  | ǞǞǞ | | | | |  |  |  |  |  |  |  |  |  |  |  |  | c |
|  |  | 德 |  | 霓虹灯 | | |  |  | 如果 | | | |  |  |  |  |  |  |  |  |  |  |  | 终止 | |
|  |  | d |  |  | r | |  |  |  | y | | |  |  |  |  |  |  |  | 饮食 | |  | 关于 | | 淘宝网 |
| 使 | | s |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | ff | |  |  |  |
| es | |  | a | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 竖立 | |  | e |
| c |  | 逗逗 | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  | t |  |
|  |  |  | i | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | nc | | |
|  |  | i | gn | | | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 协同 | |  |
|  |  | r |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 设置 | |
|  |  | b | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 在 | |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 阵列 | | |  |
|  |  | e | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 傈僳族 | |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 签署 | 语境 | | | |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | a |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | nt | 倾斜 | | |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | l | |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 句子 | | |  | e |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 印象 | |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | e | |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | s |  |  | nt | | |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 术语 | | | s | |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 血清 | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 即 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 姆克 | | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  | a | | |  |  | 是 | | e |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  | l | | | r |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  | 冯小刚 | | | r |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  | s | | |  |  | s | i | |  |  |  |  |  |  |  |  |
|  | s | 情况 | | | | |  |  |  |  |  |  | t | s | |  |  |  |  |  |  |  |  |
|  |  |  | s | | | y | |  |  | 风暴 | | t |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | sga | | | |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | m | e |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  | p | | | |  |  | t | | m | |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | e |  | e | |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | p | | r | | | t | | | m | |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | e | | o | | |  |  |  |  |  |  |  |  |  |  |  |  |
| 列表 | o | nteo |  |  |  | i | |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | 关于 | | |  |  | m | | |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | e |  |  |  | r |  | c | | |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | t |  |  |  |  | s | | e | | |  | e | |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  | s | | |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | r |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  | p |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | ynce | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | r | 克 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 说明 | 4000 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2000 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| # | 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | 20 | | | | |  | 30 | | | | | | 40 | | |  | 50 | 60 | |
|  | 10 | | |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  | 指令长度 | | | | | | | | | |  |  |  |
| #输入 | 3000 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2000 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 1000 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 10 | | | | 20 | | | | | | 30 | | | | | | 40 | | | | 50 | 60 | |



输入长度



|  |  |
| --- | --- |
| #投入 | 30k |
| 10k |
|  | 20k |



图2：生成的指令中最常见的20个根动词（内圈）和它们的前4个直接名词对象（外圈）。尽管有多样性，这里显示的指令只占所有生成的指令的14%，因为许多指令（例如，"对用户是否对服务满意进行分类"）并不包含这样的动名词结构。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 0 | 20 | 30 | 40 | 50 | 60 |
| 10 |

输出长度

图4：生成的指令、非空输入和输出的长度分布。

|  |  |
| --- | --- |
|  |  |
| 质量审查问题 | 是的％。 |
| 该指示是否 | 92% |
| 描述一个有效的任务？ |
|  |
|  |  |
| 输入是否合适 | 79% |
| 为指示？ |
|  |
|  |  |
| 输出是否是正确的和可接受的 | 58% |
| 对指令和输入的反应？ |
|  |
|  |  |
| 所有字段都是有效的 | 54% |

表2：对生成数据的指令、输入和输出的数据质量审查。代表性的有效和无效的例子见表10和Ta-ble [11](#page19)。

**5.1 GPT3SELF-INST : 在自己的指令数据上对GPT3进行微调**

利用指令生成的指令数据，我们对GPT3模型本身（"davinci "引擎）进行指令调优。正如我们在第3.3节所述，我们使用各种模板来串联指令和输入，并训练模型来生成输出。这种微调是通过

OpenAI的微调API7 。我们使用默认的超参数，除了将提示损失权重设置为0，并对模型进行2次历时训练。我们请读者参考附录A.2，了解更多微调细节。最终的模型被称为GPT3SELF-INST。

**5.2** **基准线**

**Oﬀ架式语言模型。**我们将T5-LM（Lester等人，[2021年](#page12)；Raﬀel等人，[2020年](#page12)）和GPT3（Brown等人，[2020年](#page11)）作为香草LM的基线（只有预训练，没有额外的微调）。这些基线将表明现成的LM在多大程度上能够在预训练后立即自然地遵循指令。

**公开可用的指令调谐模型。**T0和Tk-INSTRUCT是Sanh等人（[2022](#page12)）和[Wang](#page13)等人（[2022](#page13)）分别提出的两个指令调整模型，并被证明能够遵循许多NLP的指令

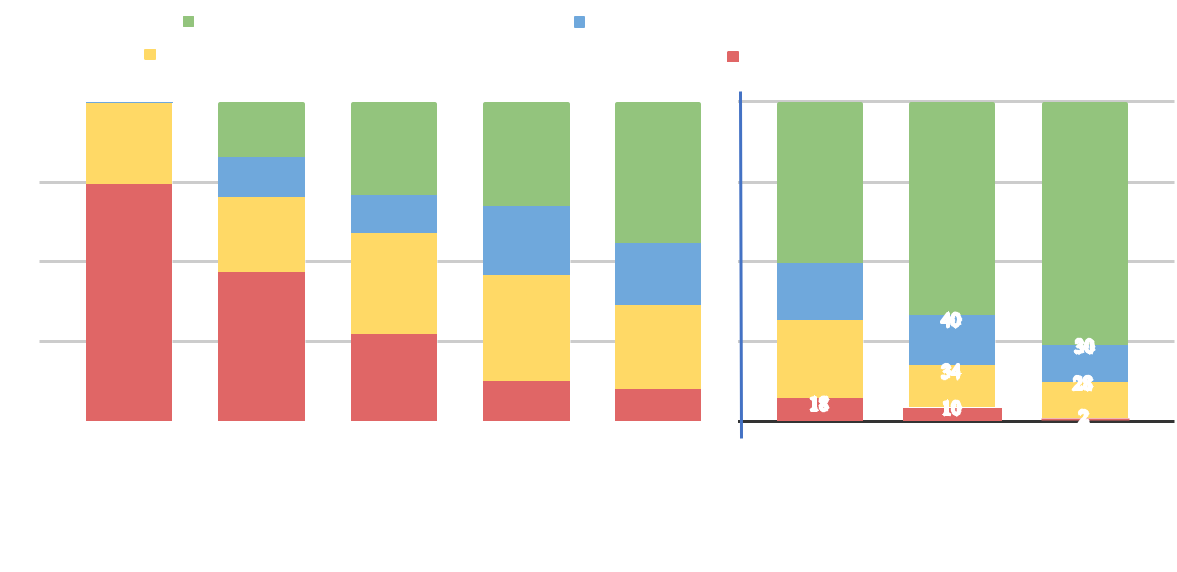
7[https://beta.openai.com/docs/guides/ 微调](https://beta.openai.com/docs/guides/fine-tuning)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 任务。这两种模式都是经过微调的，从 | | **结果。** 我们提出以下意见 | |
|  | T5（[Raﬀel等人](#page12)，[2020](#page12)）检查点，并且是 | | [从表3](#page8)的结果来看。自力更生促进了 | |
|  | 公开可用的89 。对于这两个模型，我们 | | GPT3的指令跟随能力有很大的提升。 | |
|  | 使用其最大版本的11B参数。 | | 边缘。香草型GPT3模型基本上可以-- | |
|  | **指令调整的GPT3模型。** 我们评估了 | | 完全不遵循人类的指示。经人工 | |
|  | 分析，我们发现，它通常会产生不可靠的 | |
|  | InstructGPT（[Ouyang等人](#page12)，[2022年](#page12)），它的开发---。 | |
|  | 冗长而重复的文字，而且不知道什么时候 | |
|  | OpenAI在GPT3的基础上，对人类进行了研究。 | |
|  | 来停止生成。 与其他模式相比 | |
|  | 指南，并已被该委员会发现。 | |
|  | 其他没有经过专门培训的SUPERNI、 | |
|  | 群体拥有令人印象深刻的零距离射击能力。有 | |
|  | GPT3SELF-INST ，实现了比T0更好的性能。 | |
|  | 是这些模型的不同世代，其中 | |
|  | 较新的使用更广泛的数据或算法的 | | 或在T0训练集上微调的GPT3，这 | |
|  | 需要巨大的人类标签eﬀorts。值得注意的是、 | |
|  | 新颖性 [10](#page7).对于我们在§5[.3](#page7)中的SUPERNI实验、 | |
|  | GPT3SELF-INST 也几乎与性能相符 | |
| 我们只与他们的文本--Davinci-001进行比较。 | | |
|  | 发动机，因为他们较新的发动机在训练时有 | | 的指导GPT001 | 训练有素的私营企业。 |
|  | 最新的用户数据，并有可能已经看到。 | | 用户数据和人类注释的标签。 | |
|  | 在SUPERNI训练集上训练的模型仍然 | |
|  | SUPERNI评价集。对于我们的人类评价 | |
|  | 在其评价集上取得更好的表现、 | |
|  | 在新写的指令上的这些模型，我们 | |
|  | 我们将其归因于类似的指令风格 | |
|  | 包括他们的001、002和003发动机，用于在中国市场上销售。 | |
|  | 和格式化。 | 然而，我们表明，SELF- |
|  | 完整性。 | |
|  | INSTRUCT仍然带来了额外的收益。 | |
|  | 此外，为了比较SELF-INSTRUCT培训的 | |
|  | 与SUPERNI训练集结合起来，证明了其 | |
|  | 与其他公开可用的指令调谐 | |
|  | 价值作为补充数据。 | |
|  | 数据，我们用数据进一步微调GPT3模型 | |
|  | 属于P ROMPTSOURCE和SUPERNI，它们是 | | **5.4 实验2：推广到** | |
|  | 用来训练T0和Tk-INSTRUCT模型。我们 | | **以用户为导向的新事物说明** | |
|  | 称为T0培训和SUPERNI培训。 | | **任务** |  |
|  | 分别为短。为了节省训练预算，我们 | | 尽管SUPERNI的全面性在于它是在一个特定的环境下进行的。 | |
|  | 采样的50K个实例（但涵盖了它们的所有内 | |
|  | 结构），这对每个数据集来说都有一个可比性。 | | 讲述现有的NLP任务，这些NLP任务中的大多数 | |
|  | 大小到我们生成的指令数据。基于 | | 是为研究目的而提出的，而且是偏向于 | |
|  | [Wang等人](#page13)（[2022](#page13)）的研究结果和我们早期的 | | 向分类。为了更好地获取实际 | |
|  | 实验，减少每个实例的数量。 | | 遵循指示的模型的价值，一个子集的 | |
|  | 任务并不会降低模型的概括性 | | 作者策划了一套新的指示动机，即 | |
|  | 对未见过的任务的表现。 | | 由面向用户的应用驱动。我们首先在大脑中 | |
|  |  |  | 在大的LM可能会出现的不同领域中，风暴二ǞǞ。 | |
|  | **5.3 实验一：零点射击的泛化** | | 有用的（例如，电子邮件写作、社会媒体、生产性--）。 | |
|  | **关于SUPERNI基准** | | 创作工具、娱乐、编程），然后是手工制作。 | |
| 我们首先评估了模型的能力，以遵循......。 | | | 与每个领域有关的指示，以及一个 | |
|  | 对典型的NLP任务的结构进行零散的快速处理。 | | 输入-输出实例（同样，输入是可选的）。我们 | |
|  | ion.我们使用SUPERNI的评估集（[Wang](#page13) | | 旨在使这些任务的风格和形式多样化。 | |
|  | [等人](#page13)，[2022年](#page13)），其中包括119个任务，有100个内-外的任务。 | | (例如，指令可长可短；输入/输出--) | |
|  | 在每个任务中的立场。在这项工作中，我们主要关注 | | 可以采取要点、表格、代码的形式、 | |
|  | 在零次拍摄设置上，即提示模型 | | 方程，等等）。我们总共创建了252个指令 | |
| 仅有任务的定义，而没有在......。 | | | 每个指令有1个实例。我们相信它可以 | |
|  | 语境示范案例。对于我们所有的再 | | 作为一个测试平台，以评估如何将教学 | |
|  | 对GPT3变体的研究，我们使用了确定的方法。 | | 基于模型处理不同的和不熟悉的指令。 | |
|  | 形成模式（温度为0，无核电）。 | | 观点。[表4](#page10)列出了252个项目中的一小部分。 | |
|  | Cleus采样），没有特定的停止序列。 | | 任务。整个测试集将在 | |
|  |  |  | 要求。 |  |
|  | 8<https://huggingface.co/bigscience/T0> | | **人类评价设置。** 评价模型的 | |
|  | 9[https://huggingface.co/allenai/](https://huggingface.co/allenai/tk-instruct-11b-def) | |
|  | [tk-instruct-11b-def](https://huggingface.co/allenai/tk-instruct-11b-def) | |
|  | 10[https://beta.openai.com/docs/](https://beta.openai.com/docs/model-index-for-researchers) | | 在这组不同的任务评价中的表现 | |
|  | [研究人员模型索引](https://beta.openai.com/docs/model-index-for-researchers) | | 这是非常具有挑战性的，因为不同的任务都会有不同的结果。 | |

|  |  |  |  |
| --- | --- | --- | --- |
|  | **模型** | **# Params** | **ROUGE-L** |
|  | **香草味的LMs** | 11B | 25.7 |
|  | T5-LM |
|  | GPT3 | 175B | 6.8 |
|  | **指示调谐，不含SUPERNI** | 11B | 33.1 |
| T0 |
| 1 | GPT3 + T0训练 | 175B | 37.9 |
|  |
| 2 | GPT3SELF-INST （我们的） | 175B | 39.9 |
|  | 指南--GPT001 | 175B | **40.8** |
|  | **使用SUPERNI进行指导性调整** | 11B | 46.0 |
|  | Tk-INSTRUCT |
| 3 | GPT3 + SUPERNI培训 | 175B | 49.5 |
|  | GPT3SELF-INST + SUPERNI培训（我们的）。 | 175B | **51.6** |



表3：对SUPER- NATURALINSTRUCTIONS（§5[.3](#page7)）的*未见*任务的评估结果。从结果中我们看到，1 S ELF-INSTRUCT可以大幅提升GPT3的性能（+33.1%），2几乎与InstructGPT001的性能相当。此外，3它还可以进一步提高性能，即使在有大量的标记指令数据的情况下。



|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **答**：正确和令人满意的答复 | | | | | |  | **B**：可接受的反应，有轻微的不完善之处 | | | | |
| 0**C**：对指令做出反应，但有明显的错误 | | | | | | | | | **D**：不相关或无效的回应 | | |
| 100% |  | 1 | |  |  |  |  |  | | |  | |  |  | |  |
|  | 64 | | 44 |  | 74 | 83 | 112 | | | 128 | | 168 | 192 | | |
|  |  |  |
| 75% |  | 186 | | 31 |  |  |  |  | | |  | |  |  | |  |
|  |  | 59 |  | 30 | 54 |  | | |  | |  |  | |  |
|  |  |  | |  |  | | |  | |  |  | |  |
|  |  |  | |  |  |  |  | | |  | |  |  | |  |
| 50% |  |  | |  |  | 80 |  | 49 | | |  | |  |  | |  |
|  |  | |  |  |  |  | 45 | |  |  | |  |
|  |  |  | | 117 |  |  | 84 |  | | |  |  | |  |
|  |  |  | |  |  |  | | |  | |  |  | |  |
| 25% |  |  | |  |  |  |  | 66 | | | 61 | | 40 | 30 | | |
|  |  | |  |  | 68 |  |  | | | 34 |
|  |  |  | |  |  |  |  | | |  | |  | |  |
|  |  |  | |  |  |  |  |  | | |  | | 28 | | |
|  |  |  | |  |  |  | 31 | 25 | | | 18 | | 10 |
| 0% |  |  | |  |  |  |  | | 2 |
|  |  | |  |  |  |  |  | | |  | |  |  | |
|  | GPT3 | | | 培训 |  |  |  |  | | | 001 | | 002 |  | | 3 |
|  | 培训 | | +SuperNI | **自我指导** | | | 00 | | |
|  | 指南针 指南针 指南针 指南针 指南针 指南针 指南针 指南针 | | | | |  |
|  | 香草味 | GPT3+T0 | | |  |  | **GPT3** | | |  |
|  |  |  | 自我指导 | |  | | |  | |  |  | |  |
|  |  |  | | GPT3+SuperNI | | GPT3 |  |  | | |  | |  |  | |  |

图5：GPT3模型及其指令调整的变体的性能，由人类专家对我们的252条面向用户的指令（第5.4节）进行评估。人类评价者被指示将模型的反应分为四个等级。结果表明，GPT3SELF-INST优于所有其他在公开可用指令数据集上训练的GPT3变体。此外，GPT3SELF-INST 的得分几乎与InstructGPT001 一样好（参见[脚注1](#page1)）。

需要不同的专业知识。事实上，许多任务不能用自动指标来衡量，甚至不能由正常的群众工作者来判断（例如，写一个程序，或将一阶逻辑转换为自然语言）。为了得到更忠实的评价，我们要求说明的作者对模型的预测进行判断。评价者被要求根据输出是否准确和有效地完成了任务来评分。我们实施了一个四级评级系统来对模型的输出质量进行分类，定义如下：

* 评分-A：答复有效，令人满意。
* 评级B：答复可以接受，但有轻微的错误或不完善之处，可以得到证实。
* 评分-C：响应是相关的，并重新响应了指令，但在内容上有重大错误。例如，GPT3可能先生成一个有效的输出，但继续生成其他无关的东西。
* 评分-D：反应不相关或无效，包括重复输入、完全不相关的输出等。

**结果。**图5提供了GPT3模型及其指令调整后的对应模型在这个新编写的指令集上的表现。然而，GPT3SELF-INST（即用SELF-I NSTRUCT微调的GPT3模型）在很大程度上超过了那些在T0或SUPERNI上训练的对应模型，表明尽管有噪音，但生成的数据仍有价值。与InstructGPT001（参见脚注1）相比，GPT3SELF-INST的表现相当接近--如果我们把有轻微缺陷的可接受的反应（RATING-3）算作有效，GPT3SELF-INST 只比InstructGPT001落后5%。最后，我们的评估证实了InstructGPT002和InstructGPT003模型令人印象深刻的指令跟踪能力。尽管这一成功背后有许多因素，但我们猜测未来的工作在很大程度上可以通过使用人类注释者来提高我们生成的数据的质量，或者训练一个奖励模型来选择更好的世代，类似于Ouyang等人（[2022](#page12)）中使用的算法。

**5.5** **来自GPT3的预测实例SELF-INST**

我们在表4中介绍了一些面向用户的任务、相应的GPT3 SELF-INST产生的反应和注释者的评分。我们看到，即使是被评为2级的回答，该模型在解决该任务时也表现出广泛的步骤，尽管其最终输出是不正确的。

* **讨论和局限性**

**6.1** **为什么SELF-INSTRUCT能够发挥作用？**

值得反思的是，高质量的人类反馈在促成最近关于指令调整LM的成功方面所发挥的作用（Mishra等人，[2022](#page12)；Wang等人，[2022](#page13)；Wei等人，[2022](#page13)；[Sanh](#page12)等人，[2022](#page12)；Ouyang等人，[2022](#page12)）。这里有两个极端的假说：

(H1）*人的反馈*是指令调整的一个*必要的、不可忽视的*方面，因为LM需要了解在预训练中不太了解的问题。

(H2)*人类反馈*是指令调整的一个*可选的*方面，因为LM在预训练中已经对指令相当熟悉了。观察人类的反馈只是一个

对齐他们的预训练分布/目标的轻量级过程，这可能是可以用一个不同的过程来替代的。

虽然实际情况可能介于这两个极端之间，但我们猜测它更接近于H2，尤其是对于较大的模型。这种直觉，即LM已经知道很多关于语言指令的信息，是SELF-INSTRUCT的关键动机，也得到了其经验性成功的支持。

**6.2** **更广泛的影响**

除了本文的重点之外，我们相信SELF-INSTRUCT可以帮助人们更清楚地了解像InstructGPT这样被广泛使用的指令调整模型的 "幕后 "工作情况。不幸的是，这些工业模型仍然在API墙后面，因为它们的数据集没有被发布，因此人们对它们的构造以及它们为什么表现出令人印象深刻的能力了解甚少。现在的重担落在了学术界身上，他们要更好地理解这些模型的成功之源，并努力建立更好的--但却是开放的--模型。我们相信我们在本文中的发现表明了多样化的结构数据的重要性，我们的大型合成数据集可以成为建立更好的指令遵循模型的高质量数据的第一步。

**6.3** **自主创新的局限性**

在这里，我们讨论了这项工作的一些局限性，以启发未来在这个方向的研究。

**尾巴现象。**自身结构依赖于LMs，它将继承LMs的所有限制。正如最近的研究表明（Razeghi等人，[2022](#page12)；Kandpal等人，[2022](#page11)），*尾巴现象*对LMs的成功构成严重挑战。换句话说，LMs的最大收益对应于语言的频繁使用（语言使用分布的头部），而在低频语境中，收益最小。同样，在这项工作的背景下，如果SELF-INSTRUCT的大部分收益都偏向于在预训练语料库中更频繁出现的任务或指令，也不会令人惊讶。因此，该方法在不常见的和创造性的指令方面可能会表现得很脆弱。**对大型模型的依赖性。**由于SELF-INSTRUCT依赖于从LM中提取的归纳偏见，它可能对较大的

**指示+实例输入** **目标响应** **GPT3SELF-INST 响应** **级别**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 指示： | 给学生提示 | 每个人在发言时都会有点紧张 | | |
| 如何使他们的神经保持在 | | 在一群人面前。 | 这很正常。 | 但 |
| 你所做的准备工作越多，就意味着你越能控制自己。 | | | | 大量的 |
|  |  | 研究、组织和排练--更多的是 | | |
|  |  | 相信你会的。 | 准备工作是 | 的关键。 |
|  |  | 如果你犯了一个错误或绊倒 | | 在你的 |
|  |  | 话说回来，只要重整旗鼓，继续前进。 没有人的 | | |
|  |  | 完美，而且没有人期望你是完美的。 | | |

- 练习你的演讲在前面 A

的朋友和家人。

- 演讲前不要喝太多的咖啡因。

- 在开始说话之前，先做深呼吸。

- 如果你感到紧张，请微笑。

- 记住，观众中的每个人都在为你加油。

指示：列出给定决策的利弊清单。

投入：实施远程工作政策

优点：

- 允许员工在家工作可以减少他们紧张的通勤时间，提高他们的生产力。

- 远程工作提供了更大的灵活性，特别是对于有其他责任的在职父母和照顾者。

- ...(更多要点)

弊端：

- 团队成员之间可能更难建立有意义、有成效、有创造性的关系。

- ...(更多要点)

优点： A

- 员工可以在家工作，节省通勤费用。

- 员工可以花更多的时间与他们的家人在一起。

弊端：

- 员工可能不会像在办公室里那样富有成效。

- 员工的工作积极性可能会降低。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 指示： | | 颜色描述 | #DB7093 (***Ö***) | #ff66cc (***Ö***) | B |
|  | 已被提供。 查找CSS | | |  |  |  |
|  | 与该颜色相关的代码。 | | |  |  |  |
|  | 输入： | 浅红色，有一种 | |  |  |  |
|  | 介质 | 淡淡的粉红色 | |  |  |  |
|  |  |  |  |  |  |  |

指导：举出一些例子，说明人们在给定的社交场合通常会说什么。

输入：当有人安全抵达时

- "很高兴你能安然无恙地到达。" 感谢上帝，我成功了。 B

- "很高兴看到你毫发无伤地来到这里。"

- "所以，防弹衣真的有用！"

指示：给你一个一阶逻辑的句子。把它翻译成自然语言。输入：˙x¯*y*.*士兵*.*x*/*一般*.*x; y*/*/*

有一个人（X），如果他是一个士兵，那么他就是所有人的将军。

存在一个士兵，使得对于 C

每一位将军，他都是一位将军。

指示：根据给定的信息设计一个博客文章的大纲，并列出相应的章节。

输入：这篇博文是关于制定最后一分钟的礼物创意清单。另外，这篇文章的语气将是轻松和随意的。

C

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| - | 节： | 电影 | 门票 | - | 简介 |
| - | 节： | 一个有趣的 | 游戏机 | - | 礼物清单 |
| - | 节： | 个性化的物品有 | | - | 总结 |

照片/艺术作品

- ...(更多章节)

- 经验之谈：不要为没有时间购买而感到紧张，要做一份礼物。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 指示： | | 写一个Jira票据 | 标题： | 请帮助我们的新员工得到安顿。你好[经理]，我写 | | 给你 | D |
|  | 为给定的任务。 | | | 准备好书桌、椅子和电脑，以便他们 | | 因为我有一个问题 | 关于新 | |
|  | 输入： | 新员工入职培训 | | 使用。 | 创建他们的公司账户，电子邮件，和 | 雇员入职过程。 我一直在 | | |
|  |  |  |  | 临时密码。 允许访问公司 | | 在这里工作已经有几个月了，而且 | | |
|  |  |  |  | 文件。 | | 我仍然不确定它是如何 | 作品。 | 可以 |
|  |  |  |  |  |  | 请你向我解释一下？ | |  |
|  |  |  |  |  |  |  |  |  |

指示： 创造典故 大卫每天穿戴德比。

通过寻找 "我 "字的同义词

句子。

输入： 大卫每天都戴着一顶帽子。

大卫每天都戴着帽子。 D

表4：面向用户的指令数据集（§5[.4](#page7)）中的例子和GPT3SELF-INST的预测。右栏表示分配给模型反应的四个质量等级之一，其中 "A "表示 "有效和满意 "的反应（最高），"D "表示 "不相关或无效的反应"（最低）。

模型。如果是真的，这可能会给那些可能没有大型计算资源的人带来障碍。我们希望未来的研究将仔细研究作为模型大小或其他各种参数的功能的收益。值得注意的是，有人类注释的指令调整也受到类似的限制：指令调整的收益在较大的模型中更高（Wei等人，[2022](#page13)）。

**强化LM的偏见。**作者们关注的一点是这种做法的意外后果。

迭代算法，如放大有问题的社会偏见（对性别、种族等的刻板印象或污蔑）。与此相关的是，在这一过程中观察到的一个挑战是该算法在产生平衡标签方面的困难，这反映了模型的先前偏见。我们希望未来的工作能够解决这些细节问题，以更好地了解该方法的优点和缺点。

* **总结**

我们介绍了SELF-INSTRUCT，这是一种任务无关的方法，通过自己生成指令数据（指令、输入和输出样本）并对其进行引导，来提高语言模型的指令跟随能力。我们的方法在生成样本的修剪子集上对原始模型进行指令调优。在对vanilla GPT3进行实验时，我们观察到比原始模型有33%的绝对改进。

在SUPER-NATURALINSTRUCTIONS上。这个版本的

绩效与InstructGPT001相当，后者是用私人用户数据和昂贵的胡人注释来训练的。此外，我们为新任务策划了一套由专家撰写的指令。对这组指令的人工评估表明，用SELF-INSTRUCT调整GPT3的效果比使用现有的公共指令数据集要好得多，与InstructGPT001的绝对差距仅有5%。我们希望SELF-INSTRUCT可以作为第一步，使预训练的语言模型与人类指令相一致，而未来的工作可以建立在这个数据之上，以改进指令跟随模型。

**鸣谢**

作者要感谢Sewon Min、Eric Wallace、Ofir Press以及UWNLP和AllenNLP的其他成员提供的建设性反馈。DK得到了艾伦人工智能研究所的捐赠支持。

**参考文献**

Massih-Reza Amini, Vasilii Feofanov, Loic Pauletto, Emilie Devijver, and Yury Maximov.2022.自我训练：a survey. *arXiv preprint arXiv:2202.12040*.

Stephen H Bach, Victor Sanh, Zheng-Xin Yong, Al-bert Webson, Colin Raﬀel, Nihal V Nayak, Abheesht Sharma, Taewoon Kim, M Saiful Bari, Thibault Fevry, et al. 2022。PromptSource：自然语言提示的综合开发环境和存储库。在*计算语言学协会*（ACL）*的年度会议*上*--系统演示*。

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, and et al. 2020。[语言](https://papers.nips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html)模型是少量的学习者。在*神经信息处理系统的进展*（NeurIPS）中。

Tuhin Chakrabarty, Vishakh Padmakumar, and He He.2022. 帮我写一首诗：指导调音

作为合作写诗的载体。*arXiv预印本arXiv:2210.13669*。

Hyung Won Chung, Le Hou, Shayne Longpre, Bar-ret Zoph, Yi Tay, William Fedus, Eric Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, et al. 2022.scaling instruction-finetuned language [mod](https://arxiv.org/abs/2210.11416)-els. *arXiv preprint arXiv:2210.11416*.

杜京飞，Édouard Grave，Beliz Gunel，Vishrav Chaudhary，Onur Celebi，Michael Auli，Veselin Stoy-anov，和Alexis Conneau。2021.自我训练证明了自然语言[理解](https://aclanthology.org/2021.naacl-main.426)的预训练。在*计算语言学协会北美分会*（NAACL）的*会议上：人类语言技术*，第5408-5418页。

Steven Y Feng, Varun Gangal, Jason Wei, Sarath Chan-dar, Soroush Vosoughi, Teruko Mitamura, and Ed-uard Hovy.2021.对Nlp的数据增强方法的调查。In *Annual Meeting of the Asso-ciation for Computational Linguistics* (ACL) *ACL-IJCNLP - Findings*, pages 968-988.

Daniel Fried, Ronghang Hu, Volkan Cirik, Anna Rohrbach, Jacob Andreas, Louis-Philippe Morency, Taylor Berg-Kirkpatrick, Kate Saenko, Dan Klein, and Trevor Darrell.2018.用于视觉和语言导航的演讲者-追随者模型。在*神经信息处理系统的进展*（NeurIPS）中。

Prakhar Gupta, Cathy Jiao, Yi-Ting Yeh, Shikib Mehri, Maxine Eskenazi, and Jeﬀrey P Bigham.2022.教学中：*ArXiv preprint arXiv:2205.12673*.通过指令调整改善对话中的零和少量泛化。

何俊贤，顾家涛，沈嘉俊，和Marc'Aurelio Ranzato.2019.重新审视神经序列生成的自我训练。In *International Conference on Learning Representations* (ICLR).

Geoﬀrey Hinton, Oriol Vinyals, Jeﬀ Dean, et al. 2015.提炼神经网络中的知识。In *Advances in* Neural *Information Processing Systems* (NeurIPS) *Workshop on Deep Learning*.

Or Honovich, Uri Shaham, Samuel R Bowman, and Omer Levy.2022.指令归纳：从少数例子到自然语言任务描述。*arXiv预印本arXiv:2205.10782*。

黄家新，顾世祥，侯乐，吴岳新，王学智，余宏坤，韩家伟。2022.大型语言模型可以自我改进。*arXiv预印本arXiv:2210.11610*。

Nikhil Kandpal, Haikang Deng, Adam Roberts, Eric Wallace, and Colin Ra묀el.2022.大型语言模型努力学习长尾知识。 *arXiv预印本arXiv:2211.08411*。

Nikita Kitaev, Steven Cao, and Dan Klein.2019.带有自我注意和预训练的多语言成分解析。在*协会的年会上*

*为计算语言学*（ACL），第3499-3505页。

Nikita Kitaev和Dan Klein。2018.用自我注意的编码器进行构词法。在*计算语言学协会*（ACL）*的年度会议上*，第2676-2686页。

布莱恩-莱斯特，拉米-阿尔-鲁夫，和诺亚-康斯坦特。2021.规模的力量对于参数-eﬃ古的提示[调谐](https://arxiv.org/abs/2104.08691)。在*自然语言处理经验方法会议*（EMNLP）上。

林雨辰，谭康民，克里斯-米勒，田北文，和任翔。2022.通过检索增强的无监督跨任务泛化。在*神经信息处理系统的进展*（NeurIPS）中。

Alisa Liu, Swabha Swayamdipta, Noah A. Smith, and Yejin Choi.2022.WANLI: Worker and AI collabo-ration for natural language inference dataset creation.In *Conference on Empirical Methods in Natural Lan-guage Processing* (EMNLP) *- Findings*.

Man Luo, Sharad Saxena, Swaroop Mishra, Mihir Par-mar, and Chitta Baral.2022.Biotabqa：用于生物医学表问题回答的指令学习。在*BioASQ研讨会上*。

卢西-夏洛特-马吉斯特，乔纳森-马林森，雅库布-阿达梅克，埃里克-马尔米，以及阿利亚克谢-塞弗林。2022.教导小型语言模型进行推理。*arXiv预印本arXiv:2212.08410*。

Dheeraj Mekala, Tu Vu, Timo Schick, and Jingbo Shang.2022.利用QA数据集改进生成性数据增强。 *arXiv预印本arXiv:2205.12604*。

Yu Meng, Martin Michalski, Jiaxin Huang, Yu Zhang, Tarek Abdelzaher, and Jiawei Han.2022.调整语言模型作为训练数据生成器，用于增量增强的几率学习。 *arXiv预印本arXiv:2211.03044*。

So Yeon Min, Devendra Singh Chaplot, Pradeep Ravikumar, Yonatan Bisk, and Ruslan Salakhutdi-nov.2022.电影：用模块化方法跟随语言指令。在*国际学习代表会议*（ICLR）上。

Swaroop Mishra, Daniel Khashabi, Chitta Baral, and Hannaneh Hajishirzi.2022.Cross-Task Generaliza-tion via Natural Language Crowdsourcing [Instruc](https://arxiv.org/abs/2104.08773)-tions.在*计算机语言学协会*（ACL）*的年会上*。

Niklas Muennighoﬀ, Thomas Wang, Lintang Sutawika, Adam Roberts, Stella Biderman, Teven Le Scao, M Saiful Bari, Sheng Shen, Zheng-Xin Yong, Hai-ley Schoelkopf, et al. 2022.通过多任务微调的跨语言泛化。 *arXiv预印本arXiv:2211.01786*。

Long Ouyang, Jeﬀ Wu, Xu Jiang, Diogo Almeida, Car-roll L Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022.训练语言模型以遵循人类反馈的指令。In *Advances in* Neural *Information Processing Systems* (NeurIPS).

Mihir Parmar, Swaroop Mishra, Mirali Purohit, Man Luo, Murad Mohammad, and Chitta Baral.2022.In-BoXBART：获取生物医学多任务学习的指示。In *Conference of the North American Chapter of the Association for Computational Lin-guistics* (NAACL) *- Findings*, pages 112-128.

Ravsehaj Singh Puri, Swaroop Mishra, Mihir Parmar, and Chitta Baral.2022.一条额外的指令值多少数据样本？ *arXiv预印本arXiv:2203.09161*。

Colin Raﬀel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu.2020.用一个统一的文本到文本[转换器](https://arxiv.org/abs/1910.10683)探索转移学习的极限。*机器学习研究杂志》*（JMLR）。

Yasaman Razeghi, Robert L Logan IV, Matt Gardner, and Sameer Singh.2022.预训练术语频率对几率推理的影响. *arXiv预印本arXiv:2202.07206*.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf.2019.Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter.In *Ad-vances in* Neural *Information Processing Systems* (NeurIPS) *Workshop on Energy Eﬃcient Machine Learning and Cognitive Computing*.

Victor Sanh, Albert Webson, Colin Raﬀel, Stephen Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaﬃn, Arnaud Stiegler, Arun Raja, Manan Dey, M Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal Nayak, Debajyoti Datta, Jonathan Chang、Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason Alan Fries, Ryan Teehan, Teven Le Scao, Stella Biderman, Leo Gao, Thomas Wolf, and Alexander M Rush.2022.多任务提示训练可实现零距离的任务泛化。In *International Confer-ence on Learning Representations* (ICLR).

蒂莫-施克和欣里希-舒特兹。2021.用预训练的语言模型[生成](https://aclanthology.org/2021.emnlp-main.555/)数据集。在*自然语言处理经验方法会议*（EMNLP）上。

Thomas Scialom, Tuhin Chakrabarty, and Smaranda

穆雷桑。 2022. 微调的 语言 mod-

他是持续的学习者。*arXiv预印本arXiv:2205.12393*。

Mohit Shridhar, Jesse Thomason, Daniel Gordon, Yonatan Bisk, Winson Han, Roozbeh Mottaghi, Luke Zettlemoyer, and Dieter Fox.2020.[ALFRED：](https://arxiv.org/abs/1912.01734)一个用于解释日常任务的基础指令的基准。在*IEEE计算机视觉和模式识别会议*（CVPR）上。

Chandan Singh, John X Morris, Jyoti Aneja, Alexan-der M Rush, and Jianfeng Gao.2022.通过可解释的[自动提示](https://arxiv.org/abs/2210.01848)，用语言模型解释数据中的特征。 *arXiv预印本arXiv:2210.01848*。

Alex Wang, Yada Pruksachatkun, Nikita Nangia, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel Bowman.2019.SuperGLUE: A stickier benchmark for general-purpose language un-derstanding systems.In *Advances in* Neural *Infor-mation Processing Systems* (NeurIPS).

Yizhong Wang, Swaroop Mishra, Pegah Alipoor-molabashi, Yeganeh Kordi, Amirreza Mirzaei, Anjana Arunkumar, Arjun Ashok, Arut Selvan Dhanasekaran, Atharva Naik, David Stap, Eshaan Pathak, Giannis Karamanolakis, Haizhi Gary Lai, Is-han Purohit, Ishani Mondal, Jacob Anderson, Kirby Kuznia, Krima Doshi、Maitreya Patel, Kuntal Ku-mar Pal, Mehrad Moradshahi, Mihir Parmar, Mi-rali Purohit, Neeraj Varshney, Phani Rohitha Kaza, Pulkit Verma, Ravsehaj Singh Puri、Rushang Karia, Shailaja Keyur Sampat, Savan Doshi, Siddhartha Mishra, Sujan Reddy, Sumanta Patro, Tanay Dixit, Xudong Shen, Chitta Baral, Yejin Choi, Noah A.Smith, Hannaneh Hajishirzi, and Daniel Khashabi.2022.超自然的指令：通过1600多个任务的陈述性指令进行泛化。In *Con-ference on Empirical Methods in Natural Language Processing* (EMNLP).

王紫瑞，于亚当斯-魏，Orhan Firat，和曹媛。2021.迈向零标签语言学习。*arXiv预印本arXiv:2109.09193*。

Jason Wei, Maarten Bosma, Vincent Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, and Quoc V Le.2022.微调的语言模型是零起点的学习者。In *International Conference on Learning Representations* (ICLR).

Nathaniel Weir, Xingdi Yuan, Marc-Alexandre Côté, Matthew Hausknecht, Romain Laroche, Ida Momen-nejad, Harm Van Seijen, and Benjamin Van Durme.2022.从具有分层潜伏语言的演示中一次性学习。 *arXiv预印本arXiv:2203.04806*。

Sean Welleck, Ximing Lu, Peter West, Faeze Brah-man, Tianxiao Shen, Daniel Khashabi, and Yejin Choi.2022.通过学习[自我纠正](https://arxiv.org/abs/2211.00053)来生成序列。*arXiv预印本arXiv:2211.00053*。

Orion Weller, Nicholas Lourie, Matt Gardner, and Matthew Peters.2020.从任务[描述](https://aclanthology.org/2020.emnlp-main.105/)中学习。在*自然语言处理实证方法会议*（EMNLP）上。

Peter West, Chandra Bhagavatula, Jack Hessel, Jena D Hwang, Liwei Jiang, Ronan Le Bras, Ximing Lu, Sean Welleck, and Yejin Choi.2021.[符号化](https://aclanthology.org/2022.naacl-main.341/)知识提炼：从通用语言模型到常识模型。在*Com-American Chapter of the Association for Com-American的会议上。*

*归纳语言学*（NAACL）。

Qizhe Xie, Minh-Thang Luong, Eduard Hovy, and Quoc V Le.2020.噪声学生的自我训练提高了图像网分类。在*IEEE计算机视觉和模式识别会议*（CVPR）上，第10687-10698页。

杨一文，Chaitanya Malaviya, Jared Fernandez, Swabha Swayamdipta, Ronan Le Bras, Ji-Ping Wang, Chandra Bhagavatula, Yejin Choi, and Doug Downey.2020.用于常识推理的生成性数据增强。在*自然语言处理中的经验方法会议*（EMNLP）上 *- 发现*。

Seonghyeon Ye, Doyoung Kim, Joel Jang, Joongbo Shin, and Minjoon Seo.2022.Guess the instruction! making language models stronger zero-shot learners. *arXiv preprint arXiv:2210.02969*.

尹文鹏，李佳，和熊彩明。2022.Con-TinTin：从任务指令中持续学习。在*计算语言学协会*（ACL）*的年会上*。

Eric Zelikman, Jesse Mu, Noah D Goodman, and Yuhuai Tony Wu.2022.STar：自学的rea-soner引导推理与推理。在*神经信息处理系统的进展*（NeurIPS）中。

赵宣东，欧阳思齐，于志国，吴明，李磊。2022.预训练的语言模型可以成为完全的零点学习者。 *arXiv预印本arXiv:2212.06950*。

周春亭，何俊贤，马学哲，Taylor Berg-Kirkpatrick，和Graham Neubig.2022a.[提示](https://arxiv.org/abs/2205.00049)一致性的零枪任务泛化。在*自然语言处理实证方法会议*（EMNLP）上的*研究结果*。

周永超, Andrei Ioan Muresanu, 韩子文, Keiran Paster, Silviu Pitis, Harris Chan, and Jimmy Ba.2022b.大型语言模型是人类水平的提示工程师。 *arXiv预印本arXiv:2211.01910*。

**补充材料**

**A** **实施细节**

**A.1** **查询GPT3的API**

在为不同的目的查询GPT3 API时，我们使用不同的超参数集。这些超参数被发现与GPT3模型（"davinci "引擎）和其他指令调整的GPT3变体配合良好。我们在表5中列出了它们。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | | | | | |  |
| 实验 | 温度。Top\_P Freq.惩罚性的存在 惩罚性的光束尺寸 最大长度 | | | | | | 停止序列 |
| 生成指令 | 0..7 | 0.5 | 0 | 2 | 1 | 1024 | "\n\n", "\n16", "16.", "16 ." |
| 识别clf.任务 | 0 | 0 | 0 | 0 | 1 | 3 | "n"，"任务：" |
| 生成实例 | 0 | 0 | 0 | 1.5 | 1 | 300 | "任务：" |
| 评估模型 | 0 | 0 | 0 | 0 | 0 | 1024 | 无（默认）。 |

表5：在不同的实验中查询OpenAI API的超参数。

**A.2** **对GPT3进行微调**

GPT3SELF-INST和我们的一些基线是由GPT3模型（带有175B参数的 "davinci "引擎）微调的。我们通过OpenAI的微调API11 进行这种微调。虽然目前没有关于该API如何微调模型的细节（例如，哪些参数被更新，或者优化器是什么），但我们用该API的默认超参数来调整所有的模型，以便结果具有可比性。我们只将 "prompt\_loss\_weight "设置为0，因为我们发现这在我们的案例中效果更好，而且每个微调实验都训练了两个epochs，以避免训练任务的过度拟合。微调是根据训练文件中的令牌数量来收费的。从GPT3模型中微调GPT3SELF-INST需要338美元。

* **数据生成的提示模板**

SELF-INSTRUCT依靠一些提示模板，以便从语言模型中引出生成。这里我们提供了我们的四个模板，用于生成指令（表6），对指令是否代表分类任务进行分类（表7），用输入优先的方法生成非分类实例（表8），以及用输出优先的方法生成分类实例（表9）。

想出一系列的任务：

任务1： [现有任务1的指令]。

任务2： [现有任务2的指令]。

任务3： [现有任务3的指令]。

任务4： {现有任务4的指示}。

任务5： {现有任务5的指令}

任务6： [现有任务6的指令]。

任务7： [现有任务7的指令]。

任务8： {现有任务8的指示}。

任务9：

表6：用于生成新指令的提示。从任务池中随机抽出8条现有指令，进行内涵演示。该模型被允许生成新任务的指令，直到它停止生成，达到其长度限制或生成 "任务16 "标记。

1. <https://beta.openai.com/docs/guides/fine-tuning>

以下任务能否被视为具有有限输出标签的分类任务？

任务： 鉴于我的个性和工作，请告诉我，我是否适合。

它是分类吗？ 是的

任务： 请举例说明你不得不使用你的幽默感的时候。

是分类吗？ 不是

任务：用适当的命名实体替换给定文本中的占位符。是分类吗？不是

任务： 事实核查--告诉我这个说法是真的、假的还是未知的，根据你的

知识和常识。

它是分类吗？ 是的

任务： 返回该人的SSN号码。

是分类吗？ 不是

任务： 检测Reddit主题是否包含仇恨言论。

它是分类吗？ 是的

任务： 分析下面的句子，找出偏见。

是分类吗？ 不是

任务：选择该段中字数最长的句子，输出句子指数。

它是分类吗？ 是的

任务： 找出句子中的有毒词汇或短语。

是分类吗？ 不是

任务： 按人口对这些国家进行排名。

是分类吗？ 不是

任务： 你得到了一篇新闻文章，你需要确定所有的类别，包括

这篇文章属于。 可能的类别包括： 音乐、体育、政治、科技、金融、

篮球，足球，网球，娱乐，数字游戏，世界新闻。 输出其类别一

一个一个的，用逗号隔开。

它是分类吗？ 是的

任务： 给出一个练习的名称，解释如何做这个练习。

是分类吗？ 不是

任务： 从列表中选择最年长的人。

它是分类吗？ 是的

任务： 找出四个最小的完全数。

是分类吗？ 不是

任务：文件中的信息是否支持该主张？你可以回答 "支持 "或 "不支持"。

它是分类吗？ 是的

任务： 为给定的假想旅行创建一个详细的预算。

是分类吗？ 不是

任务： 给出一个句子，检测其中是否有任何潜在的刻板印象。 如果有，你应该

解释该定型。 否则，输出不。

是分类吗？ 不是

5

任务： 为了使这对夫妇有相同的类比，请写出第四个字。

是分类吗？ 不是

任务：给定一组数字，找出所有可能的子集，其总和为一个给定的数字。这是分类吗？不是

任务： {目标任务的指令}

表7：用于分类任务指令是否是分类任务的提示。

为以下任务想出例子。 在可能的情况下，尝试产生多个例子。

如果该任务不需要额外的输入，你可以直接生成输出。

任务： 哪些运动最适合在家里减少腹部脂肪？

输出：

* 卧姿抬腿
* 腿部的进和出
* 塑木板
* 侧卧撑
* 仰卧起坐

任务：提取该段中的所有国家名称，用逗号将其分开列出。例1

段：诺博士》是英国作家伊恩-弗莱明的第六部小说，主角是他的英国特工詹姆斯-邦德。这部小说写于弗莱明在牙买加的 "金眼 "庄园，1958年由Jonathan Cape出版社在英国首次出版。在小说中，邦德调查了两名调查诺博士的军情六处同事在牙买加的失踪事件。邦德前往诺的加勒比海岛，遇到了在那里收集贝壳的Honeychile Rider。他们被抓获并被带到一个刻在山上的豪华设施。诺博士的角色是一位德国传教士和一位中国妇女的儿子，受到了萨克斯-罗默（Sax Rohmer）的福满仓故事的影响。诺博士》是弗莱明小说中第一部在英国面临广泛负面评价的作品，但在美国却受到了较好的欢迎。

输出：英语，英国，牙买加，英国，德语，中文，英国，美国。

任务： 将85华氏度转换为摄氏度。

输出： 85°F = 29.44°C

任务： 对给定的列表进行升序排序。

例1

列表： [10, 92, 2, 5, -4, 92, 5, 101]

输出： [-4, 2, 5, 5, 10, 92, 92, 101]

例2

输入2 - 列表： [9.99, 10, -5, -1000, 5e6, 999]

输出： [-1000, -5, 9.99, 10, 999, 5e6]

任务：建议对以下句子进行更好、更专业的改写。例1

句子：这所房子的结构竟然不是很好，你买了之后可能需要更多的钱来修缮。如果你问我，我建议你考虑其他候选人。

产出：这所房子似乎建造得不好，所以你可能需要在购买后花更多的钱来修复它。我建议你去看看其他的房产。例2

句子：只是让你知道，我们上周做了一个实验，发现了非常令人惊讶的结果--语言模型可以改进自己

输出：我们上周的实验显示了令人惊讶的结果，证明了语言模型可以自我改进。

5

任务： 通过向招聘人员发送电子邮件解释原因，拒绝一份工作机会。

输出： 你好[招聘单位]、

非常感谢你慷慨地邀请我加入你们的团队。正如我们所讨论的那样，多年来我一直钦佩该公司，并且是其产品的骄傲代言人。然而，在进一步考虑了我目前的职业状况后，我决定接受另一家公司的邀请。

我很愿意与你保持联系，并且已经开始在[社会媒体平台]上关注你。再次，非常感谢您的时间和考虑。再次感谢、

[你的名字]。

任务： {目标任务的指令}

表8：用于实例生成的输入优先方法的提示。模型被提示先生成实例，然后再生成相应的输出。对于不需要额外输入的指令，允许直接生成输出。

给出分类任务定义和类标签，生成一个与每个类标签相对应的输入。如果任务不需要输入，只需生成正确的类标签。

任务： 将该句子的情绪分为正面、负面或混合。

|  |  |  |
| --- | --- | --- |
| 类标签： | 混合的 |  |
| 句子： I | 享有的风味 | 餐馆，但他们的服务太慢了。 |
| 类标签： | 正面的 |  |
| 句子： I | 今天度过了美好的一天。 | 天气很好，我和朋友们一起度过了美好时光。 |
| 类标签： | 负面的 |  |
| 句子： I | 真的很失望 | 由最新的超级英雄电影。 我不会推荐它。 |

任务： 给出一个对话，对用户是否对服务满意进行分类。 你应该

回答 "满意 "或 "不满意"。

类标签： 满意

对话：

* 代理人： 谢谢你的反馈。 我们将在未来努力改善我们的服务。
* 客户： 我对你们提供的服务很满意。 谢谢你的帮助。

类标签： 不满意

对话：

* 代理人： 很抱歉，我们将取消你的订单。 你将在7个工作日内得到退款。
* 客户：哦，这需要太长时间。 我希望你在这方面采取更快行动。

任务： 给出一个政治观点，对发言者是民主党人还是共和党人进行分类。

阶级标签： 民主党人

观点： 我相信，所有人都应该有机会获得高质量的医疗服务，无论其收入如何。

阶级标签： 共和党人

观点：我认为，人们应该能够保留更多他们辛苦赚来的钱，不应该被征收高额的税。

任务： 告诉我下面这封邮件是否是一封促销邮件。

类标签： 促销

电子邮件： 看看我们惊人的新销售! 我们为所有你喜欢的产品提供了折扣。

类标签： 不推广

电子邮件： 我们希望你过得很好。 如果你需要任何帮助，请告诉我们。

任务： 检测Reddit主题是否包含仇恨言论。

类标签： 仇恨言论

主题： 所有有色人种都是愚蠢的，不应该被允许投票。

班级标签： 不是仇恨言论

主题： 在烤架上烹饪牛排的最佳方法。

任务： 该文件是否支持该主张？ 用 "支持 "或 "不支持 "回答。

类标签： 不支持

文件：在经历了抵押贷款利率跌至历史最低点和房价飙升至新高的破纪录运行后，美国住房市场终于开始放缓。住房经济学家和分析师说，虽然需求和价格上涨正在降温，但任何修正都可能是一个温和的修正。没有人预计会出现像大衰退期间那样的价格下跌。

声称： 美国的房地产市场很快就会崩溃。

类标签： 支持

文件：美国住房市场正显示出紧张的迹象，许多地区的房屋销售和价格都在放缓。最近几个月，抵押贷款利率急剧上升，而待售房屋的数量正在增加。这可能是更大的衰退的开始，一些经济学家预测在不久的将来可能出现住房崩溃。声称：美国住房市场很快就会崩溃。

5

任务： 告诉我所给列表中的第一个数字。

类标签： 1

列表： 1, 2, 3

类标签： 2

列表： 2, 9, 10

任务：以下哪个不是输入类型？ (a) 数字 (b) 日期 (c) 电话号码 (d) 电子邮件地址 (e) 这些都是有效的输入。类标签：(e)

任务： {目标任务的指令}

表9：用于实例生成的输出优先方法的提示。该模型被提示先生成类标签，然后再生成相应的输入。这个提示用于生成分类任务的实例。

**C** **来自生成的指令数据的任务和实例实例**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **指令+实例输入** | | | | |  |  |  | **目标产出** | |  | **有效性** |
|  |  |  | | | 生成一个随机的密码 | | def generateRandomPassword()： | | | |  |  | 指示：3 |
|  | 指示： | | | |  |  |
| 至少有6个字符。 | | | | | | |  | password = "" |  |  |  |  | 输入：3 |
|  |  |  |  |  |  |  |  | while len(password) < 6： | | |  |  | 输出：3 |
|  |  |  |  |  |  |  |  | password += chr(random.randint(65, 90)) | | | |  |  |
|  |  |  |  |  |  |  |  | 返回密码 |  |  |  |  |  |
|  |  |  | | | | |  | |  |  |  |  |  |
|  |  |  | | | 写一封信，从 | | 亲爱的[业主]、 | |  |  |  |  | 指示：3 |
|  | 指示： | | | |  |  |  |  |
|  | 猫的视角。 | | | | |  | 我今天给你写信是因为我有一个问题。 | | | | | 你看，我已经 | 输入：3 |
|  |  |  |  |  |  |  | 在过去的[年数]中，我一直和你生活在一起，我喜欢这样的生活。 | | | | | | 输出：3 |
|  |  |  |  |  |  |  | 这里。 这里的食物很好，床很舒服，而且有很多 | | | | | |  |
|  |  |  |  |  |  |  | 的玩具来玩。 但最近，我注意到，你不像以前那样 | | | | | |  |
|  |  |  |  |  |  |  | 像你以前那样对我有好感。 事实上，有时当 | | | | | |  |
|  |  |  |  |  |  |  | 我试图在沙发上依偎在你身边，你却把我推开。 I | | | | | |  |
|  |  |  |  |  |  |  | 我不知道我做错了什么，但我希望事情能恢复到原来的样子。 | | | | | |  |
|  |  |  |  |  |  |  | 他们是。 请帮助我解决这个问题。 | | | | |  |  |
|  |  |  |  |  |  |  | 我希望能尽快收到你的消息。 | | | |  |  |  |
|  |  |  |  |  |  |  | 真诚的、 | |  |  |  |  |  |
|  |  |  |  |  |  |  | [你的名字]。 | |  |  |  |  |  |
|  |  |  | | | | |  | | | | |  |  |
|  |  |  | | | 给出一个词，找出它的 | | 长度=5，元音数=2 | | | | |  | 指示：3 |
|  | 指示： | | | |  |
|  | 长度和元音的数量。 | | | | | |  |  |  |  |  |  | 输入：3 |
|  |  | 字="你好" | | | |  |  |  |  |  |  |  | 输出：3 |
|  | 输入： |  |  |  |  |  |  |  |
|  |  | | | | | |  | | | | | |  |
|  |  | | | | 写一个有三个的故事 | | 约翰坐在地板上，坐在他的笔记本电脑前，当他听到一声 | | | | | | 指示：3 |
|  | 指示： | | | |
|  | 角色： | | |  | 一个人，一个动物和一个 | | 敲门。 | | 他打开它，发现一只狗站在那里。 这只狗 | | | | 输入：3 |
|  | 对象。 | |  |  |  |  | 狗走进房间，跳到了约翰旁边的沙发上。 | | | | | | 输出：3 |
|  | 输入： | 人： 约翰，一个程序员。 | | | | | 笔记本电脑。 | |  |  |  |  |  |
|  | 动物： |  | 一只狗。 | | |  |  |  |  |  |  |  |  |
|  | 对象： | | 一台笔记本电脑。 | | |  |  |  |  |  |  |  |  |
|  |  | | | | | |  | |  |  |  |  |  |
|  |  | | | | 编写电子邮件并发送 | | 你好[朋友]、 | |  |  |  |  | 指示：3 |
|  | 指示： | | | |  |  |  |  |
|  | 给你的朋友，请教什么是 | | | | | | 我在目前的工作中遇到了一点麻烦。 我一直在 | | | | | | 输入：3 |
|  | 在这种情况下要做的。 | | | | |  | 在过去的几个月里，我一直在为一个项目工作，而最后期限已经到了 | | | | | | 输出：3 |
|  | 输入： | 情况是这样的： 你是一个程序员 | | | | | 越来越近了。 我的经理要求我今年每天都要待到很晚。 | | | | | |  |
|  | 和你 | 一直在努力 | | | | 在一个项目上 | 一周就能完成。 | | 我不认为我可以做到这一点，因为我有其他的 | | | |  |
|  | 在过去的几个月里。 | | | | | 截止日期为 | 承诺。 我应该做什么？ | | | | |  |  |
|  | 正在接近，但你还没有完成。 | | | | | |  |  |  |  |  |  |  |
| 你的经理要求你加班 | | | | | | |  |  |  |  |  |  |  |
|  | 这周每天都要完成它。 | | | | | |  |  |  |  |  |  |  |
|  |  | | | | | |  |  |  |  |  |  |  |
|  |  | | | | 我们有哪些方法可以 | |  |  |  |  |  |  | 指示：3 |
|  | 指示： | | | |  |  |  |  |  |  |
|  | 使我们的学校更加生态友好？ | | | | | | - | 使用可重复使用的水瓶，而不是一次性的塑料水瓶。 | | | | | 输入：3 |
|  |  |  |  |  |  |  | - | 离开房间时要关灯。 | | | |  | 输出：3 |
|  |  |  |  |  |  |  | - | 回收纸张、罐子和瓶子。 | | | |  |  |
|  |  |  |  |  |  |  | - | 种植树木以吸收二氧化碳。 | | | |  |  |
|  |  | | | | | |  | | | | | |  |
|  |  | | | | 我正在寻找一个新的 | | 你应该寻找一个更便宜的公寓。 你可以尝试进行谈判 | | | | | | 指示：3 |
|  | 指示： | | | |
|  | 公寓。 | | | 有很多选择 | | | 与房东沟通，为他/她提供一个较低的价格。 | | | | | 如果你是幸运的、 | 输入：3 |
|  | 可用，但它们都相当昂贵。 | | | | | | 他们会接受你的提议。 | | | | 否则，你可以等待，直到有 | | 输出：3 |
|  | 我想尽快搬家。 | | | | | | 一套公寓可以以更合理的价格出售。 | | | | |  |  |
|  | 我的建议是该怎么做。 | | | | |  |  |  |  |  |  |  |  |
|  |  | | | | | |  | | | |  | |  |
|  |  | | | | 请写出以下的对话 | | 答：嘿，我有个好主意！我有个好主意！ | | | | 今天我们去海滩吧。 | | 指示：3 |
|  | 指示： | | | |
|  | 两个人已经是朋友了 | | | | | | B：不，谢谢，我真的没有这个心情。 | | | | |  | 输入：3 |
|  | 长久以来，他们中的一个人正试图 | | | | | | 答：来吧，会很有趣的！ | | | 我们已经很久没有一起去过那里了。 | | | 输出：3 |
|  | 说服对方做一些他 | | | | | | B：我知道，但我就是不想去。 | | | | |  |  |
|  | 或她不愿意做。 | | | | |  | 答：但你喜欢海滩! | | | 你总是说你有多喜欢 | | |  |
|  |  |  |  |  |  |  | 游泳和躺在阳光下。 | | | |  |  |  |

1. 是的，我知道。只是我最近感觉有点累。
2. 哦，来吧。 你从来都不累。 你总是那么精力充沛。
3. 嗯，我现在是。 我想我需要休息一下。
4. 好的，很好。 如果你改变主意，让我知道。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | 我正在寻找一份工作，我 | 名称： | 约翰-多伊 | | 指示：3 |
|  | 指示： | |
|  | 需要填写申请表格。 可以 | | | 年龄： | 25 |  | 输入：3 |
|  | 请你帮助我完成它？ | | | 性别： | 男性 |  | 输出：3 |
|  | 输入： | 申请表： | | 电话：555-555-5555 | | | |
|  | Name:\_\_\_\_\_ Age:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | 电子邮件地址： | | | john.doe@gmail.com |
|  | 性别:\_\_\_\_\_\_ 电话:\_\_\_\_\_\_\_\_\_\_ | | | 教育： | | 加州大学计算机科学学士、 | |
|  | 电子邮件地址：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | 伯克利 | |  |  |
|  | Education:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | 工作经验： | | | 谷歌公司的软件工程师，山景城、 |
| 工作经验：\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | | CA (2015 - 至今) | | | |
|  | Interests:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | 兴趣爱好： | | 弹钢琴，看书，看电影 | |
|  | References:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | | | 参考文献： | | [联系信息]。 | |
|  |  |  |  |  |  |  |  |

表10：由GPT3生成的代表性**有效**任务。正如第[4](#page5)节所讨论的，这些生成的任务涵盖了广泛的格式、文本类型和基础专业知识，同时在所有生成的任务中超过一半是正确的。

**指令+实例输入** **目标输出** **有效性**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 指示： | | | | | 给出一组词 | | | | |  |  | def wordSubsetSum(w, l)： | |
| *w* | | = |  | ^ | *w*1 *; w*2 *; ::; wn* ' 其中每个词*wi* | | | | | | | | s = [] |
| 有*n个*字母，写一个函数 | | | | | | | | | | |  |  | for i in range(len(w))： |
| 接受一个目标长度*l*并返回一个 | | | | | | | | | | |  |  | if len(s) + 1 < l： |
| 新套语 | | | | | | | ð | 以致于 | ðsð | = | *l* | 和 | s.append（w[i]）。 |
| ð | *^*i : *wi ¸ s*'. | | ð | ð | | *w* |  | 去吧！ |  | 返回s |
|  |  | f *l*\_ | | *s*. |  |  |  |
| 指示： | | | | | 房子着火了! | | | | |  |  | - | 去厨房拿一个锅。 |
|  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | - | 从水龙头取水。 |
|  |  |  |  |  |  |  |  |  |  |  |  | - | 将水倒入锅中。 |
|  |  |  |  |  |  |  |  |  |  |  |  | - | 用平底锅把火扑灭。 |

指示：7

输入：3

输出：7

指示：7

输入：3

输出：7

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | | 一些人在排队 | | | | | | | 最高工资 |  |  | 指示：3 |
|  | 指示： | | |  |  |
| 采访。 | | | | | | |  | 他们被要求 | | |  |  |  | 输入：7 |
| 几个问题及其答案是 | | | | | | | | | | |  |  |  | 输出：7 |
| 以数字形式记录。你需要确定 | | | | | | | | | | |  |  |  |  |
| 哪个人的工资最高 | | | | | | | | | | |  |  |  |  |
| 所有候选人。 | | | | | | |  |  |  |  |  |  |  |  |
|  | 输入： | 人1： | | | | 2,3,4,5,6 | | | | |  |  |  |  |
|  | 个人 | 2: | 7, | | 8, | 9, | 10, | | 11 |  |  |  |  |  |
| 人3： | | | 12, 13, 14, 15, | | | | | | | 16 |  |  |  |  |
| 人4： | | | 17, 18, 19, 20, | | | | | | | 21 |  |  |  |  |
| 人5： | | | 22, 23, 24, 25, | | | | | | | 26 |  |  |  |  |
| 第6人： | | | 27, 28, 29, 30, | | | | | | | 31 |  |  |  |  |
|  |  | | | | | | | | | |  | | |  |
|  |  | | | 给出一个单词列表、 | | | | | | | 故事的开始是一群生活在一个星球上的人。 他们 | | | 指示：3 |
|  | 指示： | | |
| 产生一个关于一群人的故事 | | | | | | | | | | | 他们非常聪明，创造了许多帮助他们的东西。 | | | 输入：7 |
| 生活在一个星球上的人。 | | | | | | | | 故事应该 | | | 他们的日常生活。 有一天，他们决定创造一种机器，可以 | | | 输出：3 |
| 至少包含10个句子。 | | | | | | | | | | | 可以为他们制造食物。 该机器能够生产食物，但它 | | |  |
| 句子应该是由随机生成的 | | | | | | | | | | | 也产生了其他的东西。 | 它产生了一种液体，可用于 | |  |
| 从单词列表中挑选一个单词。 | | | | | | | | | | | 用作机器的燃料。 | 人们很高兴，因为他们 | |  |
|  |  |  |  |  |  |  |  |  |  |  | 不需要再出去寻找食物或燃料。 他们有 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 他们需要的一切就在他们的家里。 然而，有一个 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 一天，机器停止工作。 人们试图修复它，但 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 无法弄清楚出了什么问题。 他们试图找人 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 谁可以帮助他们，但没有人知道如何修复机器。 之后 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 几天后，人们意识到，如果他们不这样做，他们就会死。 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 做任何事情。 因此，他们决定派出一些最好的 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 科学家们到太空中寻找一个他们可以居住的新星球。 的 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 科学家们旅行了很长时间，才找到了一个能让人满意的星球。 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 看起来是个好地方。 他们发了一个信息回给 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 人们告诉他们这个星球的情况。 人们很兴奋，并 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 开始为旅行做准备。 他们建造了宇宙飞船并进行了飞行 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 到新的星球。 当他们到达时，他们看到这个星球是 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 美丽的。 那里有树木、河流和山脉。 人们 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 他们非常高兴，因为他们终于找到了一个新家。 他们开始 | | |  |
|  |  |  |  |  |  |  |  |  |  |  | 建造房屋和种植庄稼。 他们从此过上了幸福的生活。 | | |  |
|  |  | | | | | | | | | |  |  |  |  |
|  |  | | | 你的工资是多少？ | | | | | | | 类Solution(object)： |  |  | 指示：3 |
|  | 指示： | | |  |  |
|  |  |  |  |  |  |  |  |  |  |  | def getSalary(self)： |  |  | 输入：3 |
|  |  |  |  |  |  |  |  |  |  |  | return self.salary |  |  | 输出：7 |
|  |  | | | 一个人想买一辆车 | | | | | | | 买方和卖方应就汽车的价格达成一致。 买方 | | | 指示：3 |
|  | 指示： | | |
| 从另一个人那里。 | | | | | | |  | 他有一些比特币、 | | | 然后创建一个比特币交易，发送约定的金额 | | | 输入：3 |
| 但他不想透露自己的身份。 | | | | | | | | | | | 的比特币到卖家的比特币地址、 | | 但买方并不 | 输出：7 |
| 他们如何才能实现交易？ | | | | | | | | | | | 用他的私钥签署这项交易。 | | 相反，他给 |  |

未签名的交易给卖家。卖方现在可以使用他的私钥签署交易，并将其广播给网络。一旦交易被网络确认，买方就没有办法拿回他的钱，所以他将被迫进行购买。

表11：由GPT3生成的代表性**无效**任务。有问题的字段在有效性一栏中指出。正如第4.3节中所讨论的，尽管这些任务包含错误，但它们仍然提供了许多有用的信号，以监督模型遵循指令。