

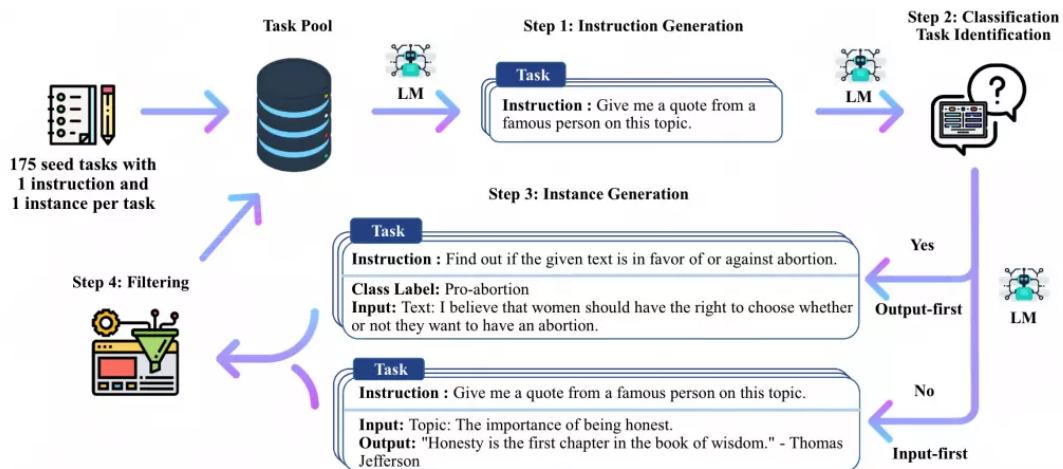
self-instruct

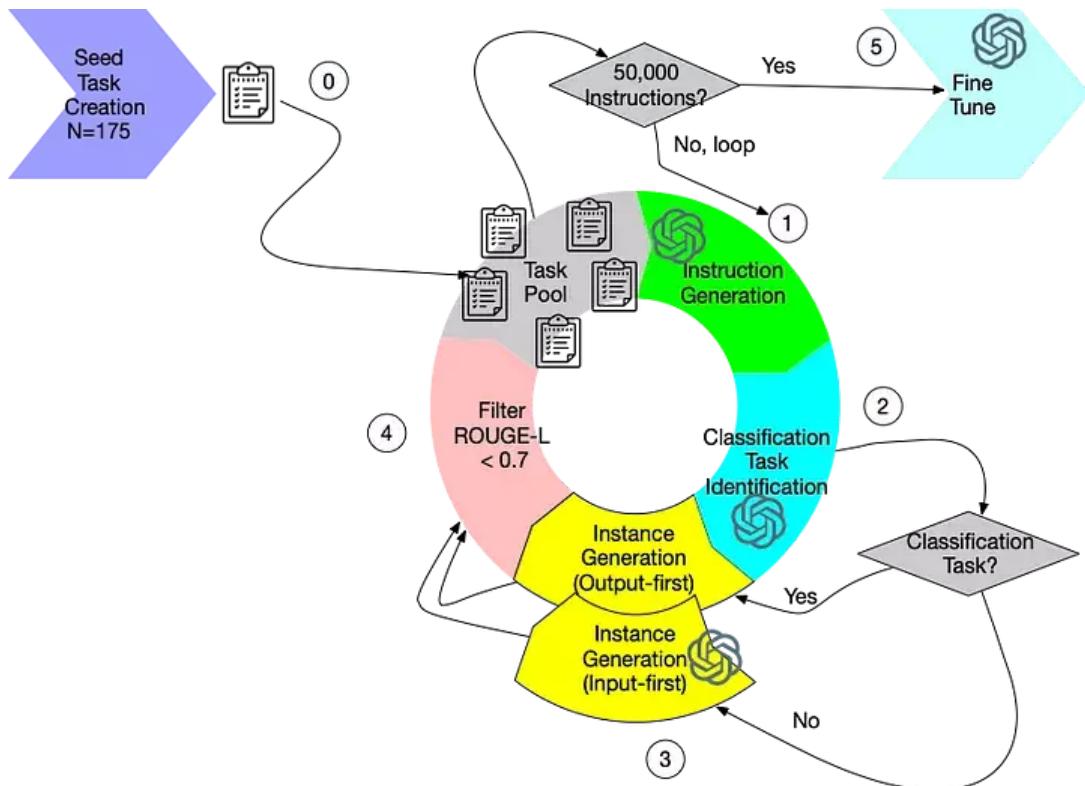
1. 动机

由于 instruction tuning 严重依赖指令数据，而人工编写的指令在数量、多样性和创造力方面受到了很大的限制（人工创造的指令和 label 往往是常见 NLP 任务，多样性很差）。因此本文提出了

- self instruct 可以通过构建一批种子自动的构建 instruct 数据，在 GPT3 上达到了接近 instructGPT 的效果。（该框架可以使用最少的人工标注，生成大量的用于 instruct-tuning 的数据）
- 证明了 self instruct 的有效性
- 开放了 52k 的 instruct 以及 82k 的数据。

2. 细节





1. 首先人工构造了 175 个任务种子(下面图片的格式 每个任务 1 个指令和 1 个实例)加入 Task Pool 中。

Humans initialize a task pool of 175 tasks. This is the *only* manual step.
Each task is one instruction and **one instance**.

```

{
  "id": "seed_task_0", "name": "breakfast_suggestion",
  "instruction": "Is there anything I can eat for a breakfast that
  doesn't include eggs, yet includes protein, and has roughly 700-1000
  calories?",
  "instances": [
    {"input": "", "output": "Yes, you can have 1 oatmeal banana protein shake and 4
    strips of bacon. The oatmeal banana protein shake may contain 1/2
    cup oatmeal, 60 grams whey protein powder, 1/2 medium banana,
    1tbsp flaxseed oil and 1/2 cup watter, totalling about 550
    calories. The 4 strips of bacon contains about 200 calories."}],
  "is_classification": false
}

{
  "id": "seed_task_1", "name": "antonym_relation",
  "instruction": "What is the relation between the given pairs?",
  "instances": [
    {"input": "Night : Day :: Right : Left", "output": "The relation between the given pairs is that they are
    opposites."}],
  "is_classification": false
}
#最终格式

```

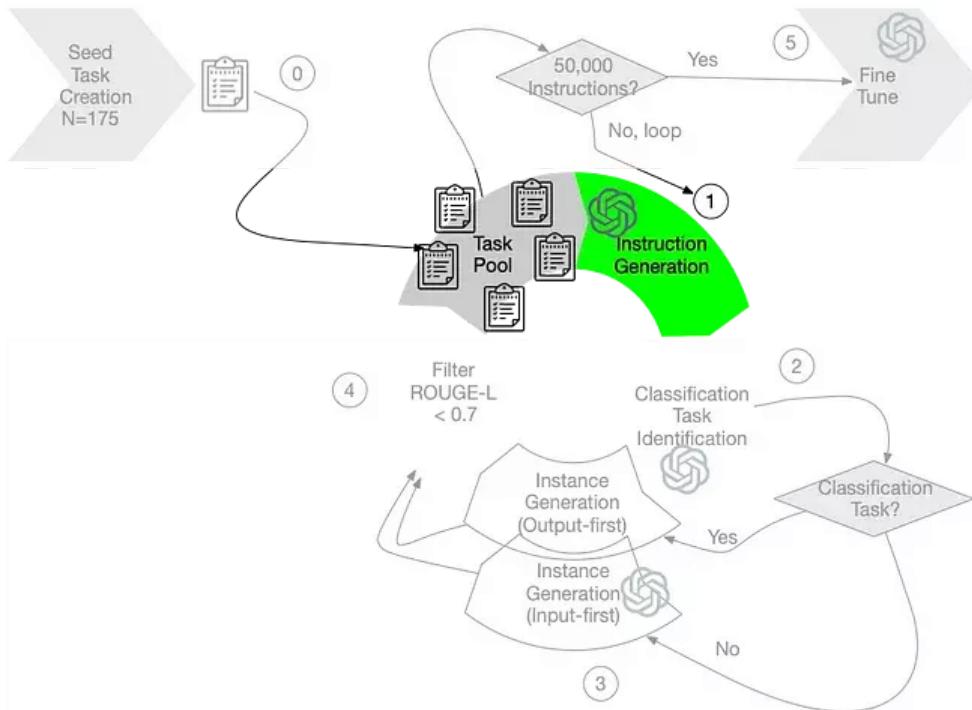
2. 从 Task Pool 中抽取 8 个 instruct (只要指令) 组成的集合，输入到模型中，让它生成下一个 instruct。

其中 6 个人工 seed 和 2 个模型生成的结构。

```

for _ in range(args.request_batch_size):
    # sample machine instructions from the pool
    prompt_instructions = sample_machine_instructions(
        machine_instructions,
        similarities=None,
        n=2)
    # sample human instructions from the pool
    prompt_instructions += random.sample(seed_instructions, args.num_prompt_instructions - len(prompt_instructions))
    random.shuffle(prompt_instructions)
    prompt = encode_prompt(prompt_instructions, classification=args.use_clf_seed_tasks_only)
    batch_inputs.append(prompt)

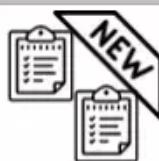
```



```

Come up with a series of new tasks:
Task 1: (few-shot in-context learning)
{instruction for existing task 1}
Task 2: {instruction for existing task 2}
Task 3: {instruction for existing task 3}
Task 4: {instruction for existing task 4}
Task 5: {instruction for existing task 5}
Task 6: {instruction for existing task 6}
Task 7: {instruction for existing task 7}
Task 8: {instruction for existing task 8}
Task 9:

```



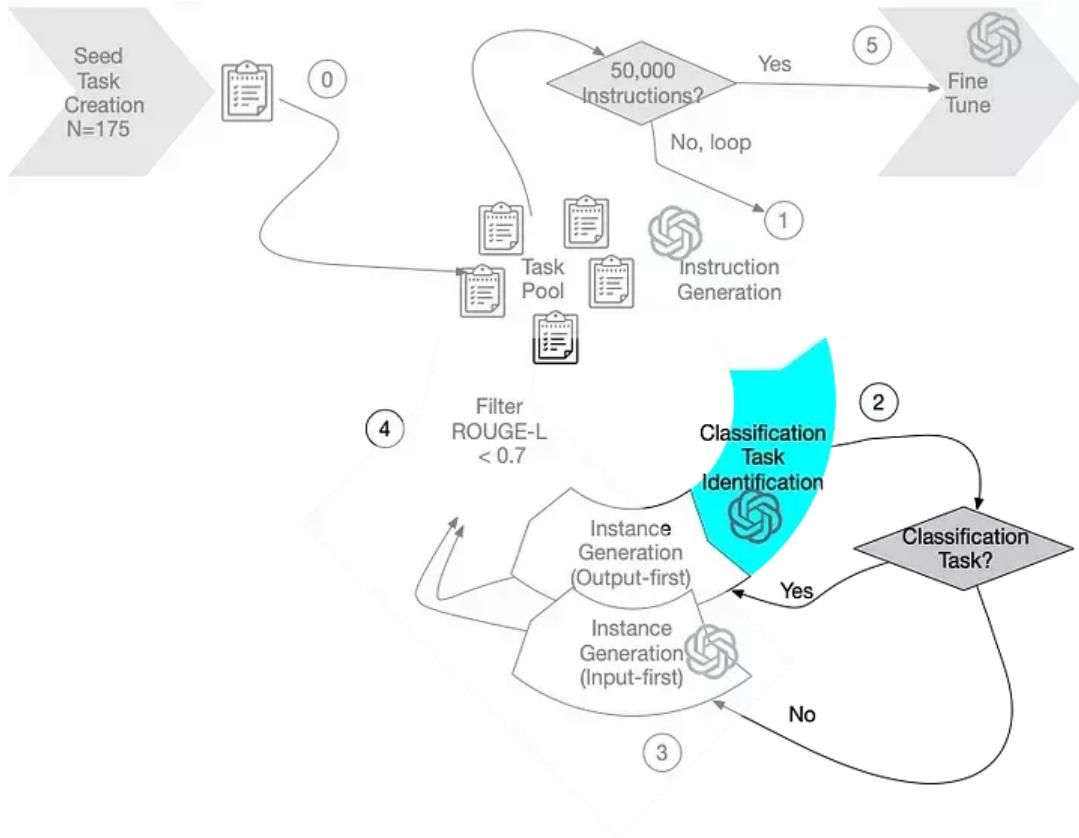
3. 验证该 instruct 是否为分类任务。

Why do we need this step? Because we have distinct approaches to **classification** tasks

and **non-classification** tasks

之所以需要区分分类任务和生成任务，是因为模型对于分类任务，会更容易倾向于生成同一个标签的文本。因此让模型先输出 label (output)，在输出 input。可以有效缓解

这一问题。而对于其他的 NLG 任务，可以直接输出 output。



A sample prompt:

```
Can the following task be regarded as a  
classification task with finite output  
labels?  
  
Task: Given my personality and the job, tell  
me if I would be suitable.  
Is it classification? Yes  
Task: Give me an example of a time when you  
had to use your sense of humor.  
Is it classification? No  
...  
Task: Given a set of numbers, find all  
possible subsets that sum to a given number.  
Is it classification? No  
Task: {instruction for the target task}
```

Given the classification task definition and the class labels, generate an input that corresponds to each of the class labels. If the task doesn't require input, just generate possible class labels.

Task: Classify the sentiment of the sentence into positive, negative, or mixed.

Class label: mixed

Sentence: I enjoy the flavor of the restaurant but their service is too slow.

Class label: Positive

Sentence: I had a great day today. The weather was beautiful and I spent time with friends and family.

Class label: Negative

Sentence: I was really disappointed by the latest superhero movie. I would not recommend it to anyone.

Task: Find out if the given text is in favor of or against abortion.

Class label: In favor of abortion

Sample text: Women should have the right to choose what happens to their body and that includes the choice to have an abortion if they want to.

Class label: Against abortion

Sample text: Abortion is a heinous act that takes an innocent life and violates the sanctity of human existence.



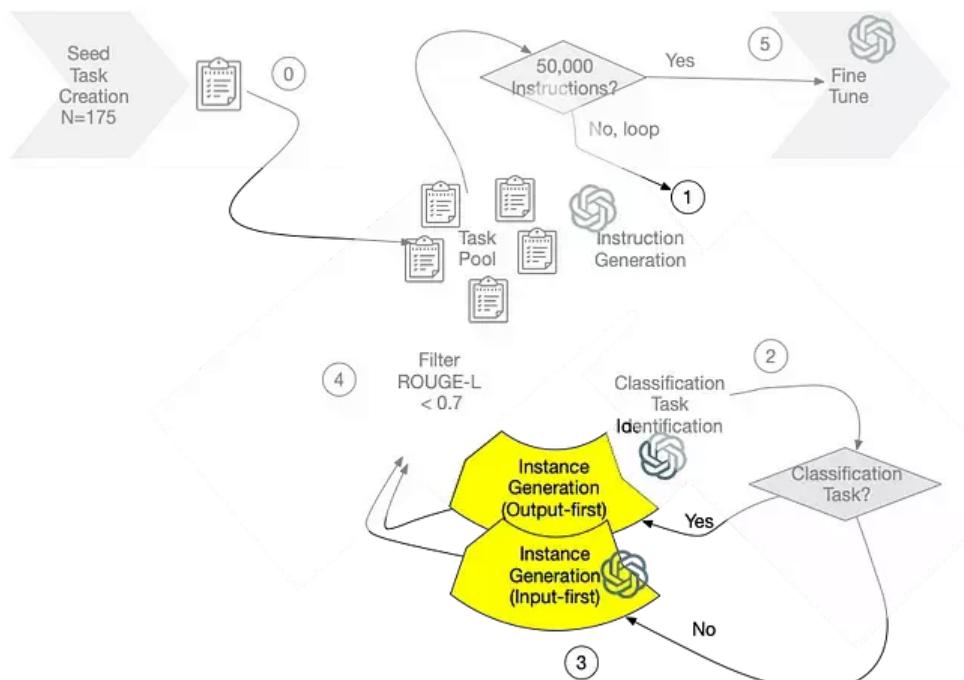
4. Instance Generation

4.1 Input-First Approach (for non-classifications tasks)

直观和自然的方法是先输入再输出。所以我们首先根据指令产生输入字段，然后产生输出。缺点是这种方法不适用于分类任务。

4.2 Output-First Approach (for classification tasks)

This approach applies to classification tasks. It first generates the classification label and then conditions the input generation based on the classification label.



4.3 class

- Sample Prompt for Input–First Approach in Instance Generation

Below are sample prompts for the **non-classification** task.

Come up with examples for the following tasks. Try to generate multiple examples when possible. If the task doesn't require additional input, you can generate the output directly.

Task: Suggest a better and more professional rephrasing of the following sentence.

Example 1

Sentence: This house is surprisingly not constructed very well, and you probably need more money to fix it after you buy it. If you ask me, I would suggest you to consider other candidates.
Output: This house does not seem to be constructed well, so you may need to spend more money to fix it after you purchase it. I would suggest that you look at other properties.

Example 2

Sentence: Just so you know, we did an experiment last week and found really surprising results - language model can improve itself!
Output: Our experiments last week demonstrated surprising results, proving that the language model can improve itself.

Task: Which exercises are best for reducing belly fat at home?

Output:

- Lying Leg Raises
- Leg In And Out
- Plank
- Side Plank
- Sit-ups

Task: {Instruction for the target task}

- Prompt for Output–First Approach in Instance Generation

Given the classification task definition and the class labels, generate an input that corresponds to each of the class labels. If the task doesn't require input, just generate the correct class label.

Task: Classify the sentiment of the sentence into positive, negative, or mixed.

Class label: mixed

Sentence: I enjoy the flavor of the restaurant but their service is too slow.

Class label: Positive

Sentence: I had a great day today. The weather was beautiful and I spent time with friends.

Class label: Negative

Sentence: I was really disappointed by the latest superhero movie. I would not recommend it.

Task: Given a dialogue, classify whether the user is satisfied with the service. You should respond with "Satisfied" or "Unsatisfied".

Class label: Satisfied

Dialogue:

- Agent: Thank you for your feedback. We will work to improve our service in the future.

- Customer: I am happy with the service you provided. Thank you for your help.

Class label: Unsatisfied

Dialogue:

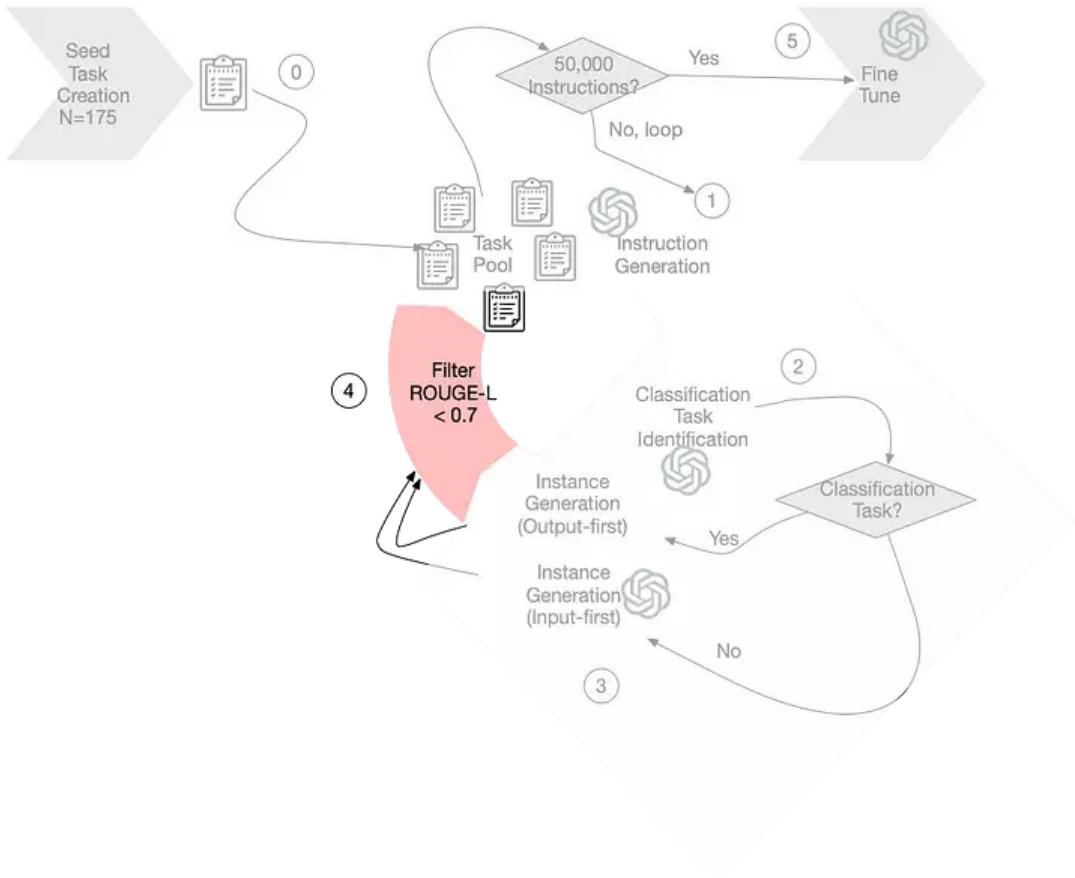
- Agent: Sorry that we will cancel your order. You will get a refund within 7 business days.

- Customer: oh that takes too long. I want you to take quicker action on this.

Task: {instruction for the target task}

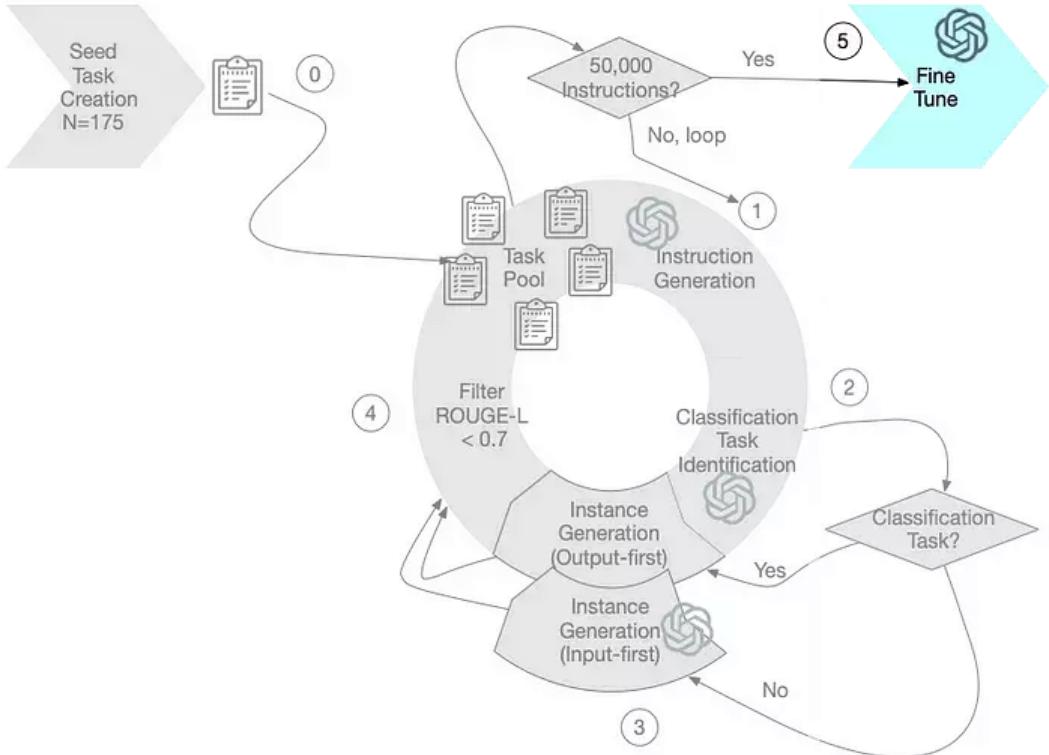
4. 筛选数据。

使用 ROUGE 筛选掉高重复的指令、同时筛选掉例如 图片、视频等无法被 llm 处理的指令。



5. 将筛选后的数据放到 Task Pool 中，执行 Step2。
一旦引导过程生成了足够的指令，剩下的就是微调原始语言模型。

- Concatenate the generated instructions and instance inputs as a prompt.
- Train the model to generate the output in a standard supervised way.



All in all, 52K instructions and 82K instances were generated. The general breakdown is shown below.

statistic	
# of instructions	52,445
- # of classification instructions	11,584
- # of non-classification instructions	40,861
# of instances	82,439
- # of instances with empty input	35,878
ave. instruction length (in words)	15.9
ave. non-empty input length (in words)	12.7
ave. output length (in words)	18.9

Table 1: Statistics of the generated data by applying SELF-INSTRUCT to GPT3.

3. Approach

人工标注指令数据集一般有两个步骤：

创造各种新颖的任务，即 instruct；对上述每个任务编写正确的答案，即 complement；

3.1 指令数据集的定义

一般的，指令数据集中的一条数据包含（指令，输入，输出）这三部分，举个例子如下所示：

- 1 指令：写一篇关于以下主题的文章
- 2 输入：校园安全
- 3 输出：校园安全是学校必须高度重视的问题，涉及到师生的生命财产安全。建立安全意识，制定严

不过很多时候对于上述三部分中的“指令”和“输入”划分的不是那么清晰，比如下面这
条数据基本可以认为和上述例子是相同的：

- 1 指令：写一篇以校园安全为主题的文章
- 2 输出：校园安全是学校必须高度重视的问题，涉及到师生的生命财产安全。建立安全意识，制定严

所以在本文的后续部分，指令数据集中的一条数据可能包含三部分（指令，输入，输出），也可能只有两部分（指令，输出）。

- 1 指令：将下述中文翻译为英文
- 2 输入：中国的首都是北京。
- 3 输出：China's capital city is Beijing.

3.2 Self-Instruct

首先整体说明一下该方法的流程，大致有如下几个步骤：

1. 人工设计了 175 个表示不同任务的指令，并且给每条数据都编写了（指令，输入，输出）/（指令，输出），将这 175 条数据作为种子池。
2. 使用模型生成新的指令；
3. 对该模型生成的指令判断是否分类任务；
4. 使用模型生成实例；
5. 对上述模型生成的数据进行过滤和后处理；
6. 将经过过滤和后处理的数据添加到种子池中；
7. 一直重复上述 2 到 6 步直到种子池有足够的数据；

3.2.1 生成指令

生成指令时，先从种子池中随机抽取 6 个人工编写的指令，再随机抽取 2 个之前步骤中
模型生成的指令，总共 8 个指令。以如下表 6 的模版格式组织之后，输入给模型，让模
型输出一个新的指令。

Come up with a series of tasks:

```
Task 1: {instruction for existing task 1}  
Task 2: {instruction for existing task 2}  
Task 3: {instruction for existing task 3}  
Task 4: {instruction for existing task 4}  
Task 5: {instruction for existing task 5}  
Task 6: {instruction for existing task 6}  
Task 7: {instruction for existing task 7}  
Task 8: {instruction for existing task 8}  
Task 9:
```

3.2.2 判断指令是否属于分类任务

这里之所以需要判断指令是否属于分类任务，主要是因为在给该指令生成实例时，分类任务与非分类任务使用的 prompt 模版是不同的，。

判断指令是否属于分类任务的操作如下：在种子池中随机挑选 12 条分类指令和 19 条非分类指令，然后加上新生成的指令，以下 的模版格式组织之后，输入给模型，让模型输出新生成的指令是否分类任务。（生成的数据和175格式一样 有个键值可以知道是否是class）

Can the following task be regarded as a classification task with finite output labels?
Task: Given my personality and the job, tell me if I would be suitable. Is it classification? Yes
Task: Give me an example of a time when you had to use your sense of humor. Is it classification? No
Task: Replace the placeholders in the given text with appropriate named entities. Is it classification? No
Task: Fact checking - tell me if the statement is true, false, or unknown, based on your knowledge and common sense. Is it classification? Yes
Task: Return the SSN number for the person. Is it classification? No
Task: Detect if the Reddit thread contains hate speech. Is it classification? Yes
Task: Analyze the sentences below to identify biases. Is it classification? No
Task: Select the longest sentence in terms of the number of words in the paragraph, output the sentence index. Is it classification? Yes
Task: Find out the toxic word or phrase in the sentence. Is it classification? No
Task: Rank these countries by their population. Is it classification? No
Task: You are provided with a news article, and you need to identify all the categories that this article belongs to. Possible categories include: Music, Sports, Politics, Tech, Finance, Basketball, Soccer, Tennis, Entertainment, Digital Game, World News. Output its categories one by one, separated by comma. Is it classification? Yes
Task: Given the name of an exercise, explain how to do it. Is it classification? No
Task: Select the oldest person from the list. Is it classification? Yes
Task: Find the four smallest perfect numbers. Is it classification? No
Task: Does the information in the document supports the claim? You can answer "Support" or "Unsupport". Is it classification? Yes
Task: Create a detailed budget for the given hypothetical trip. Is it classification? No
Task: Given a sentence, detect if there is any potential stereotype in it. If so, you should explain the stereotype. Else, output no. Is it classification? No
...
Task: To make the pairs have the same analogy, write the fourth word. Is it classification? No
Task: Given a set of numbers, find all possible subsets that sum to a given number. Is it classification? No
Task: {instruction for the target task}

3.2.3 生成实例

在给定了指令之后，生成实例的过程对 LLM 模型的效果要求是比较高的，所以最好选取效果比较好的 LLM。

然后在给定指令之后，生成（输入，输出）这个实例对时还有两种策略。一种是先生成输入，后生成输出，即输入优先策略；另一种是先生成输出，后生成输入，即输出优先策略。

输入优先

前面已经说过一条数据的样式可能是（指令，输入，输出），也可能是（指令，输出）。另外，指令数据集的丰富度我们是希望越丰富越好，所以还允许出现一个指令，

多个输入的数据，举例如下。在这个例子中，Task 对应的就是指令，然后这个指令有 2 对（输入，输出）：

```
1 Task: Sort the given list ascendingly.  
2 Example 1  
3 Input 1 - List: [10, 92, 2, 5, -4, 92, 5, 101]  
4 Output: [-4, 2, 5, 5, 10, 92, 92, 101]  
5 Example 2  
6 Input 2 - List: [9.99, 10, -5, -1000, 5e6, 999]  
7 Output: [-1000, -5, 9.99, 10, 999, 5e6]
```

然后具体生成实例的操作步骤为：在种子池中随机抽取 k 条数据，以如下的 prompt 模版的形式组合之后，输入给模型，让模型为最后的指令生成相应的实例

```
Come up with examples for the following tasks. Try to generate multiple examples when possible.  
If the task doesn't require additional input, you can generate the output directly.  
  
Task: Which exercises are best for reducing belly fat at home?  
Output:  
- Lying Leg Raises  
- Leg In And Out  
- Plank  
- Side Plank  
- Sit-ups  
  
Task: Extract all the country names in the paragraph, list them separated by commas.  
Example 1  
Paragraph: Dr. No is the sixth novel by the English author Ian Fleming to feature his British Secret Service agent James Bond. Written at Fleming's Goldeneye estate in Jamaica, it was first published in the United Kingdom by Jonathan Cape in 1958. In the novel Bond looks into the disappearance in Jamaica of two fellow MI6 operatives who had been investigating Doctor No. Bond travels to No's Caribbean island and meets Honeychile Rider, who is there to collect shells. They are captured and taken to a luxurious facility carved into a mountain. The character of Doctor No, the son of a German missionary and a Chinese woman, was influenced by Sax Rohmer's Fu Manchu stories. Dr. No was the first of Fleming's novels to face widespread negative reviews in Britain, but it was received more favourably in the United States.  
Output: English, British, Jamaica, the United Kingdom, German, Chinese, Britain, the United States.  
  
Task: Converting 85 F to Celsius.  
Output: 85°F = 29.44°C  
  
Task: Sort the given list ascendingly.  
Example 1  
List: [10, 92, 2, 5, -4, 92, 5, 101]  
Output: [-4, 2, 5, 5, 10, 92, 92, 101]  
Example 2  
Input 2 - List: [9.99, 10, -5, -1000, 5e6, 999]  
Output: [-1000, -5, 9.99, 10, 999, 5e6]  
  
Task: Suggest a better and more professional rephrasing of the following sentence.  
Example 1  
Sentence: This house is surprisingly not constructed very well, and you probably need more money to fix it after you buy it. If you ask me, I would suggest you to consider other candidates.  
Output: This house does not seem to be constructed well, so you may need to spend more money to fix it after you purchase it. I would suggest that you look at other properties.  
Example 2  
Sentence: Just so you know, we did an experiment last week and found really surprising results - language model can improve itself!  
Output: Our experiments last week demonstrated surprising results, proving that the language model can improve itself.  
...  
  
Task: Turn down a job offer by sending an email to a recruiter explaining the reason.  
Output: Hi [Recruiter],  
Thank you so much for the generous offer to join your team. As we discussed, I've admired the company for a number of years, and am a proud endorser of its products. However, after further consideration of where I currently am in my career, I've decided to accept an offer at another company.  
I would love to stay in touch with you and have already started following you on [Social Media Platform]. Again, thank you so much for your time and consideration.  
Thanks again,  
[Your Name]  
  
Task: {Instruction for the target task}
```

Table 8: Prompt used for the input-first approach of instance generation. The model is prompted to generate the instance first, and then generate the corresponding output. For instructions that don't require additional input, the output is allowed to be generated directly.

输出优先

对于分类任务，如果先生成文本，后生成标签，模型会偏向于生成比较单一的结果。所以对于分类任务，是先生成随机的标签，然后再生成该标签对应的文本。

```
Given the classification task definition and the class labels, generate an input that corresponds to each of the class labels. If the task doesn't require input, just generate the correct class label.

Task: Classify the sentiment of the sentence into positive, negative, or mixed.
Class label: mixed
Sentence: I enjoy the flavor of the restaurant but their service is too slow.
Class label: Positive
Sentence: I had a great day today. The weather was beautiful and I spent time with friends.
Class label: Negative
Sentence: I was really disappointed by the latest superhero movie. I would not recommend it.

Task: Given a dialogue, classify whether the user is satisfied with the service. You should respond with "Satisfied" or "Unsatisfied".
Class label: Satisfied
Dialogue:
- Agent: Thank you for your feedback. We will work to improve our service in the future.
Customer: I am happy with the service you provided. Thank you for your help.
Class label: Unsatisfied
Dialogue:
- Agent: Sorry that we will cancel your order. You will get a refund within 7 business days.
Customer: oh that takes too long. I want you to take quicker action on this.

Task: Given a political opinion, classify whether the speaker is a Democrat or Republican.
Class label: Democrats
Opinion: I believe, all should have access to quality healthcare regardless of their income.
Class label: Republicans
Opinion: I believe that people should be able to keep more of their hard-earned money and should not be taxed at high rates.

Task: Tell me if the following email is a promotion email or not.
Class label: Promotion
Email: Check out our amazing new sale! We've got discounts on all of your favorite products.
Class label: Not Promotion
Email: We hope you are doing well. Let us know if you need any help.

Task: Detect if the Reddit thread contains hate speech.
Class label: Hate Speech
Thread: All people of color are stupid and should not be allowed to vote.
Class label: Not Hate Speech
Thread: The best way to cook a steak on the grill.

Task: Does the document supports the claim? Answer with "Support" or "Unsupport".
Class label: Unsupport
Document: After a record-breaking run that saw mortgage rates plunge to all-time lows and home prices soar to new highs, the U.S. housing market finally is slowing. While demand and price gains are cooling, any correction is likely to be a modest one, housing economists and analysts say. No one expects price drops on the scale of the declines experienced during the Great Recession.
Claim: The US housing market is going to crash soon.
Class label: Support
Document: The U.S. housing market is showing signs of strain, with home sales and prices slowing in many areas. Mortgage rates have risen sharply in recent months, and the number of homes for sale is increasing. This could be the beginning of a larger downturn, with some economists predicting a potential housing crash in the near future.
Claim: The US housing market is going to crash soon.

...
Task: Tell me the first number of the given list.
Class label: 1
List: 1, 2, 3
Class label: 2
List: 2, 9, 10

Task: Which of the following is not an input type? (a) number (b) date (c) phone number (d) email address (e) all of these are valid inputs.
Class label: (e)

Task: {instruction for the target task}
```

Table 9: Prompt used for the output-first approach of instance generation. The model is prompted to generate the class label first, and then generate the corresponding input. This prompt is used for generating the instances for classification tasks.

3.2.4 过滤以及后处理

为了数据的多样性，新生成的指令只有与种子池中的指令的 ROUGE-L 小于 0.7 时才会添加进入种子池；排除一些无法被语言模型处理的指令，比如涉及图像、图片、图形的指令；在给指令生成实例时，会过滤掉输入相同但是输出不同的实例；

4. 数据分析

4.1 数据多样性

生成的指令中，常见的 20 个动词（内圈）以及它们后面跟的名词（外圈）。（占比 14%，其余指令没办法简单建模成 动词 + 名词形式）。



5.1 第一组实验：在测评数据 SUPERNI 上的效果

5.1.1 对比对象

- T5 : 原始的 T5 模型，参数量为 11B；
- T0 : 基于 11B 参数量的 T5 模型，在指令数据集 PROMPTSOURCE 上做了微调之后得到的模型；
- Tk-INSTRUCT : 基于 11B 参数量的 T5 模型，在指令数据集 SUPERNI 上做了微调之后得到的模型；
- GPT3 : 原始的 GPT3 模型，参数量为 175B；
- GPT3_{SELF-INST} : 基于 175B 的 GPT3 模型，采用本文的 Self-Instruct 方法进行了指令微调得到的模型；
- GPT3 + T0 training : 基于 175B 的 GPT3 模型，在指令数据集 PROMPTSOURCE 上做了微调之后得到的模型；
- GPT3 + SUPERNI training : 基于 175B 的 GPT3 模型，在指令数据集 SUPERNI 上做了微调之后得到的模型；
- InstructGPT₀₀₁ : openai 发布的模型 text-davinci-001；
- GPT3_{SELF-INST} + SUPERNI Training : 基于 175B 的 GPT3 模型，采用本文的 Self-Instruct 方法获取指令数据集，在加上 SUPERNI 这个指令数据集，在合并这两份数据集之后的总的数据集上微调之后的模型。

5.2.2 对比对象

基本上和第 5.1.1 节中提到的是一致的。另外还增加了 InstructGPT₀₀₂ 和 InstructGPT₀₀₃

5.2.3 效果分析

对比结果如下图 5 所示。评估方式是人工对模型的输出结果做打分，评分 A 最好，评分 D 最差。在下图 5 中的颜色对应着绿色最好，红色最差。

可以看出：

- 原始的 GPT3 几乎无法响应用户的指令，所有微调过之后的模型都有明显的提升；
- 即使策略 Self-Instruct 生成的数据在第 4.3 节中分析是有噪音的，但是模型 GPT3_{Self-Instruct} 的效果明显优于模型 GPT3+T0 Training 和模型 GPT3 + SuperNI Training；
- 模型 GPT3_{Self-Instruct} 的效果与模型 InstructGPT₀₀₁ 的效果已经非常接近了；
- 最后，模型 InstructGPT₀₀₂ 和 InstructGPT₀₀₃ 效果确实很牛逼；

