# Self-Instruct: Aligning Language Model with Self-Generated Instructions

Yizhong Wang♣ Yeganeh Kordi◇ Swaroop Mishra♡ Alisa Liu♣ Noah A. Smith♣+ Daniel Khashabi♠ Hannaneh Hajishirzi♣+

Why do we need a self generation method for instructions?

- Large "instruction-tuned" language medals are popular today.
- They generalize zero-shot to new tasks.
- LLM developments are powered by two key components: large pretrained LMs and human-written instruction data (e.g., PROMPT-SOURCE and SUPER-NATURAL-INSTRUCTIONS or SUPERNI)
- Human-written instructions are often limited in quantity, diversity and creativity.
- SELF-INSTRUCT: Framework for improving the instruction-following capabilities of pretrained LLMs by bootstrapping off their own generations.
  - Generates instructions, input, and output samples from a LM, then filters invalid or similar ones before using them to finetune the original model.

How does Self-Instruct work broadly?
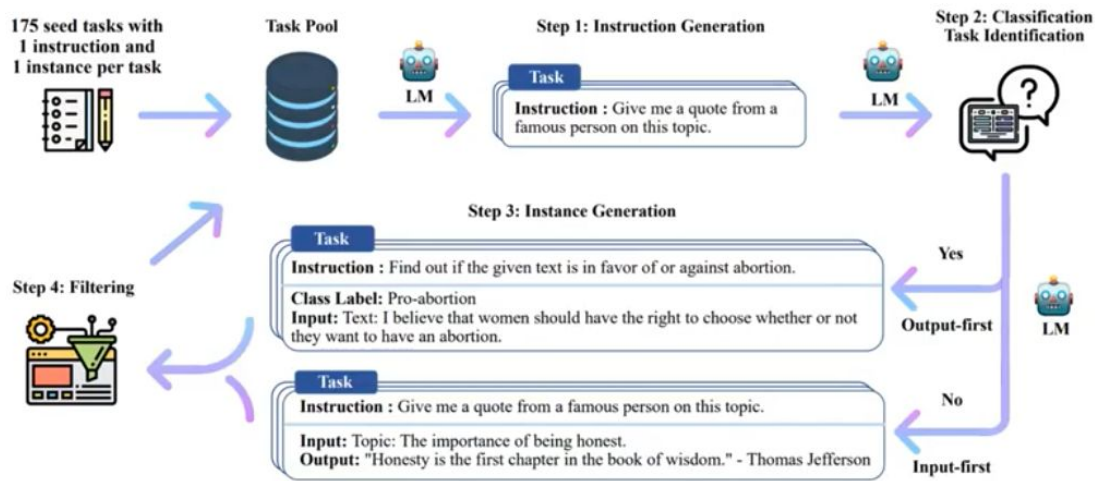
1. Generating task instructions.
2. Determining if the instruction represents a classification task.
3. Instance generation with either an input-first or output-first approach.
4. Filtering low-quality data.

How does Self-Instruct work? Instruction Generation

- Start with 175 tasks (1 instruction and 1 instance for each task).
- Prompt used for generating new instructions.
- 8 existing instructions are randomly sampled from the task pool for in-context demonstration.
- The model is allowed to generate instructions for new tasks, until it stops its generation , or reaches its length limit.
- Of the 8 instructions, 6 are from the human written tasks, and 2 are from the model-generated tasks in previous steps to promote diversity.

```
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:
```

How does Self-Instruct work? Classification Task Identification

- Prompt the LM in a few-shot way
- Using 12 classification instructions and 19 non-classification instructions from the seed tasks.

```
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:  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}
```

How does Self-Instruct work? Instance Generation

- Given the instruction and their task type, generate instance for each instruction independently.
- Model needs to
  - Understand what the target task is, based on the instruction
  - Figure out what additional input fields are needed
  - Generate input and output.
- Prompt with instruction-input-output in-context examples from other tasks.
- Input-first Approach: ask a LLM to come up with the input fields and then outputs.
  - Can generate input biased towards one label , especially for classification tasks (e.g for grammar error detection, it usually generates grammatical input).
- Output-first Approach for classification tasks: first generates the possible class label, and then condition the input generation on each class label.

```
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:  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}
```

```
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.
```

How does Self-Instruct work? Filtering and Postprocessing

- Remove Instructions if
    - ROUGH-L similarity with any existing instruction >= 0.7, for diversity.
    - Contains some specific keywords (e.g., image, picture, graph) that usually can not be processed by LLMs.
    - Too long or too short
- Filter out instances if
    - They are exactly the same or those with the same input but different outputs.
    - Instance output is a repetition of the input
- Fine tuning GPT3-SELF-INST from GPT3 on the entire generated data cost $338 (2 epochs)

Self-Instruct dataset properties

- 52K instructions and more than 82K instances.
- GPT3 LM ("davinci" engine)

Quality Review Question	Yes %
Does the instruction describe a valid task?	92%
Is the input appropriate for the instruction?	79%
Is the output a correct and acceptable response to the instruction and input?	58%
All fields are valid	54%

Instruction: Given an address and city, come up with the zip code.
Input:
Address: 123 Main Street, City: San Francisco
Output: 94105

Instruction: I am looking for a job and I need to fill out an application form. Can you please help me complete it?
Input:
Application Form:
Name:_____ Age:_____ Sex:_____
Phone Number:_____ Email Address:_____
Education:_____ ...
Output:
Name: John Doe Age: 25 Sex: Male
Phone Number: ...

Instruction: How to write a code for converting degrees fahrenheit to celsius.
Input: Null
Output:
def convert_fahrenheit_to_celsius(fahr):
    celsius = (fahr - 32) * 5 / 9
    return celsius

Instruction: Write a letter from the perspective of a cat.
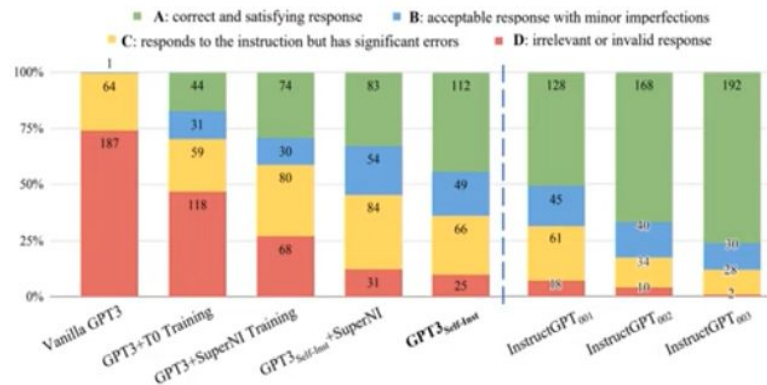**Input:** Null
Output:
Dear [Owner], I am writing to you today because I have a problem. You see, I've been living with you for the past [Number of Years] and I love it here. The food is great, the bed is comfy, and there are plenty of toys to play with. But lately, I've noticed that you're not as affectionate towards me as you used to be. In fact, sometimes when I try to cuddle up next to you on the couch, you push me away ...
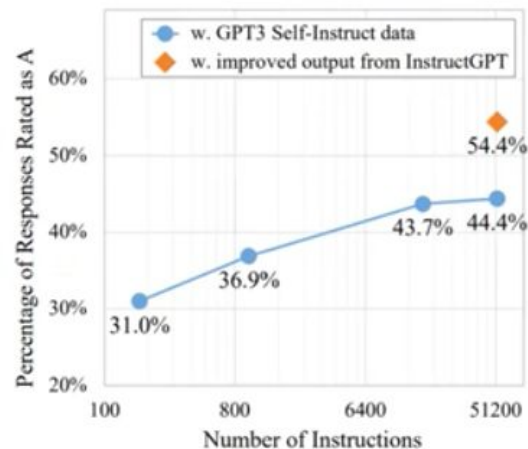
How does GPT-3 Self-Instruct perform?

- We further finetune GPT model with data from PROMPTSOURCE (T0 training) and SUPERNI.
- Zero-Shot Generalization on SUPERNI benchmark
  - SELF-INSTRUCT can boost GPT3 performance by a large margin (+33.1%)
  - Nearly matches the performance of InstructGPT001.
  - It can further improve the performance even when a large amount of labeled instruction data is present.
- Generalization to User-oriented Instructions on Novel Tasks
  - Across various domains where large LMs may be useful (e.g., email writing, social media, productivity tools, entertainment, programming) , they craft instructions related to each domain along with an input-output instance.
  - Get GPT3-Self-Instruct outputs judge by humans.
  - SELF-INSTRUCT outperforms using existing public instruction datasets by a large margin.

Model	# Params	ROUGE-L
Vanilla LMs		
T5-LM	11B	25.7
GPT3	175B	6.8
Instruction-tuned w/o SUPERNI		
T0	11B	33.1
GPT3 + T0 Training	175B	37.9
GPT3$_{SELF-INST}$ (Ours)	175B	39.9
InstructGPT$_{001}$	175B	**40.8**
Instruction-tuned w/ SUPERNI		
Tk-INSTRUCT	11B	46.0
GPT3 + SUPERNI Training	175B	49.5
GPT3$_{SELF-INST}$ + SUPERNI Training (Ours)	175B	**51.6**

- **A**: correct and satisfying response   **B**: acceptable response with minor imperfections
- **C**: responds to the instruction but has significant errors   **D**: irrelevant or invalid response

How does GPT-3 Self-Instruct perform?

Instruction + Instance Input	Target Response	GPT3_SELF-INST Response	Rating
Instruction: Give students tips on how to keep their nerves under control during class presentations.	Everyone gets a little nervous when speaking in front of a group. That's normal. But the more preparation you've done-meaning plenty of researching, organizing, and rehearsing-the more confident you'll be. Preparation is the key. And if you make a mistake or stumble over your words, just regroup and keep going. Nobody's perfect, and nobody expects you to be.	- Practice your presentation in front of friends and family. - Don't drink too much caffeine before the presentation. - Take deep breaths before you start speaking. - If you feel nervous, smile. - Remember that everyone in the audience is rooting for you.	A
Instruction: Make a list of the pros and cons of the given decision. **Input:** Implementing a Remote Working Policy	Pros: - Allowing employees to work from home can reduce their stressful commutes and increase their productivity. - Remote working provides greater flexibility, particularly for working parents and caregivers who have other responsibilities. - ...(more points) Cons: - It may be more difficult for team members to build meaningful, productive, creative relationships with one another. - ...(more points)	Pros: - Employees can work from home and save on commuting costs. - Employees can spend more time with their families. Cons: - Employees may not be as productive as they would be in the office. - Employees may not be as motivated to work.	A
Instruction: A user story is an informal, general explanation of a software feature written from the perspective of the end-user or customer. Write a user story for a given software. **Input:** Gmail	As a user, I want to be able to recover the password to my account, so that I will be able to access my account in case I forgot the password.	As a user, I want to be able to send emails to my friends and family. I want to be able to attach files to my emails. I want to be able to search for emails in my inbox.	A
Instruction: A color description has been provided. Find the CSS code associated with that color. **Input:** A light red color with a medium light shade of pink	#DB7093 (■)	#ff66cc (■)	B

- As we increase number of instructions, we see consistent improvement as we grow the data size.
- Use Instruct PT003 to re-generate the output field of all out instances given the instruction and input. We then use this improved version of our data to finetune GPT3.

10

Summery

- SELF-INSTRUCT: a method to improve the instruction-following ability of LMs via their own generation of instruction data.
- A large-scale dataset of 52K instruction for diverse tasks, and fine tuning GPT3 on this data leads to a 33% absolute improvement on SUPERNI over the original GPT3.

Thank You