



SELF-INSTRUCT: Aligning Language Model with Self Generated Instructions

Y Wang 저술 · 2022

00

Summary

- 🍌 Instruction을 LLM을 통해 생성하고 그걸로 튜닝해도(self-instruct) 성능이 좋아짐을 보여주는 연구 (GPT-3 API 사용)
- 🍌 Instruction도 Instance도 LLM로 생성하기 때문에 challenge하고 human의 개입이 잘 안 들어간 방법론을 써서 InstructGPT001 모델과 비슷한 성능을 냈다가 Contribution
- 🍌 여기서 이야기하는 Instruction은 NLP task라기보다 InstructGPT에서의 prompt중 명령 관련 표현이라고 이해하면 좋음
- 🍌 LLM에서 꺼내써야 되기 때문에 (inductive bias에 의존) LM이 커야됨

```
results = make_gpt3_requests(  
    engine="davinci",  
    prompts=batch_inputs,  
    max_tokens=1024,  
    temperature=0.7,  
    top_p=0.5,  
    frequency_penalty=0,  
    presence_penalty=2,  
    stop_sequences=["\n\n", "\n16", "16.", "16 ."],  
    logprobs=1,  
    n=1,  
    best_of=1,  
    api_key=args.api_key,  
    organization=args.organization,  
)  
  
if max(rouge_scores) > 0.7:  
    continue  
all_instructions = seed_instructions + machine_instructions  
most_similar_instructions = {  
    all_instructions[i] : rouge_scores[i] for i in np.argsort(rouge_scores)[-10:][::-1]  
}  
machine_instructions.append(inst)  
fout.write(json.dumps({  
    "instruction": inst,  
    "most_similar": most_similar_instructions,  
    "avg_similarity_score": float(np.mean(rouge_scores)),  
    "metadata": metadata,  
    "request_idx": request_idx  
}) + "\n")
```

<https://github.com/yizhongw/self-instruct>

```
{  
    "id": "seed_task_23",  
    "name": "cooking_recipe",  
    "instruction":  
        "Give me a cooking recipe for the following food."  
    ,  
    "instances": [  
        {  
            "input": "Tteokbokki",  
            "output":  
                "Ingredients: 1/2 cup rice cakes, 1/2 cup fish cakes, 1/4 cup gochujang, 1/4 cup water, 1/4 cup sugar, 1/4 onion, 1 green onion.  
                \n  
                Instructions: 1. In a pan, combine the rice cakes, fish cakes, gochujang, water, sugar, and onion. 2. Cook over medium heat for 5-7 minutes, stirring occasionally. 3. Add the green onion and cook for 1-2 minutes."  
        }  
    ],  
    "is_classification": false  
}
```

```
{  
    "id": "seed_task_163",  
    "name": "oldest_person",  
    "instruction":  
        "Select the oldest person from the list.",  
    "instances": [  
        {  
            "input":  
                "George Washington, Confucius, Michael Jordan, Michelangelo"  
            ,  
            "output": "Confucius"  
        }  
    ],  
    "is_classification": true  
}
```

00

Summary

- 🍌 Instruction을 LLM을 통해 생성하고 그걸로 튜닝해도(self-instruct) 성능이 좋아짐을 보여주는 연구 (GPT-3 API 사용)
- 🍌 Instruction도 Instance도 LLM로 생성하기 때문에 challenge하고 human의 개입이 잘 안 들어간 방법론을 써서 InstructGPT001 모델과 비슷한 성능을 냈다가 Contribution
- 🍌 여기서 이야기하는 Instruction은 NLP task라기보다 InstructGPT에서의 prompt중 명령 관련 표현이라고 이해하면 좋음
- 🍌 LLM에서 꺼내써야 되기 때문에 (inductive bias에 의존) LM이 커야됨

```
"name": "breakfast_suggestion", "instruction": "Is there  
"name": "antonym_relation", "instruction": "What is the  
"name": "one_sentence_description", "instruction": "Gene  
"name": "harmful_stereotype_example", "instruction": "De  
"name": "email_subject_generation", "instruction": "Gene  
"name": "interview_question_answering", "instruction": "  
"name": "new_year's_resolutions", "instruction": "Brains  
"name": "idiom_meaning", "instruction": "Explain the fol  
"name": "english_haiku_generation", "instruction": "Gene  
"name": "movie_recommendation", "instruction": "Recommen  
"name": "game_suggestion", "instruction": "Suggest some  
"name": "grocery_list", "instruction": "Make a grocery  
"name": "explain_behavior", "instruction": "Explain hum  
"name": "dependency_parsing", "instruction": "Dependenc
```

```
results = make_gpt3_requests(  
    engine="davinci",  
    prompts=batch_inputs,  
    max_tokens=1024,  
    temperature=0.7,  
    top_p=0.5,  
    frequency_penalty=0,  
    presence_penalty=2,  
    stop_sequences=["\n\n", "\n16", "16.", "16 ."],  
    logprobs=1,  
    n=1,  
    best_of=1,  
    api_key=args.api_key,  
    organization=args.organization,  
)  
  
if max(rouge_scores) > 0.7:  
    continue  
all_instructions = seed_instructions + machine_instructions  
most_similar_instructions = {  
    all_instructions[i] : rouge_scores[i] for i in np.argsort(rouge_scores)[-10:][::-1]  
}  
machine_instructions.append(inst)  
fout.write(json.dumps({  
    "instruction": inst,  
    "most_similar": most_similar_instructions,  
    "avg_similarity_score": float(np.mean(rouge_scores)),  
    "metadata": metadata,  
    "request_idx": request_idx  
}) + "\n")
```

<https://github.com/yizhongw/self-instruct>

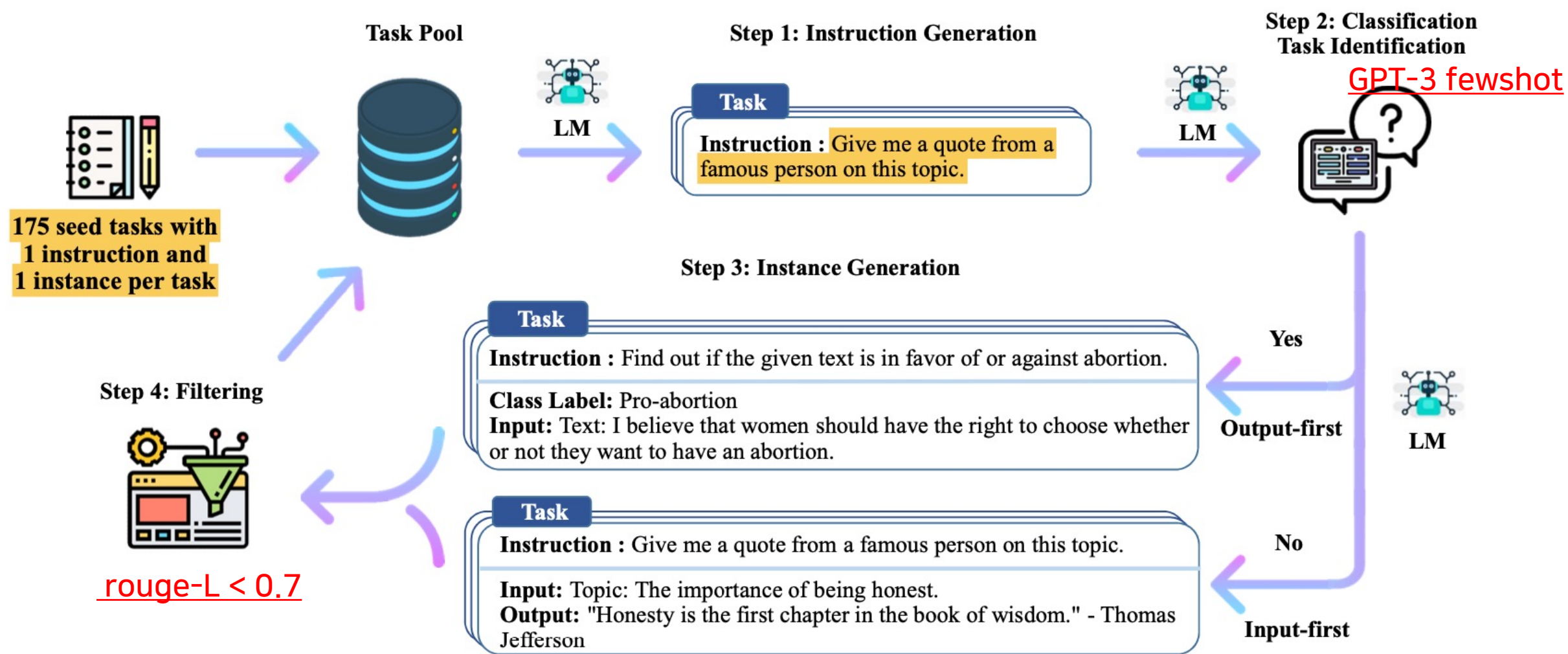
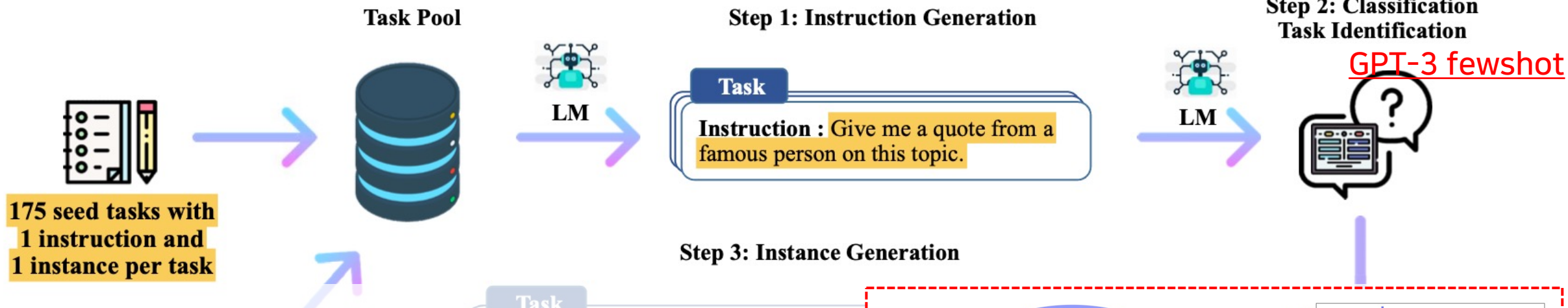


Figure 1: A high-level overview of SELF-INSTRUCT. The process starts with a small seed set of tasks (one instruction and one input-output instance for each task) as the task pool. Random tasks are sampled from the task pool, and used to prompt an off-the-shelf LM to generate both new instructions and corresponding instances, followed by filtering low-quality or similar generations, and then added back to the initial repository of tasks. The resulting data can be used for the instruction tuning of the language model itself later to follow instructions better. Tasks shown in the figure are generated by GPT3. See Table 10 for more creative examples.



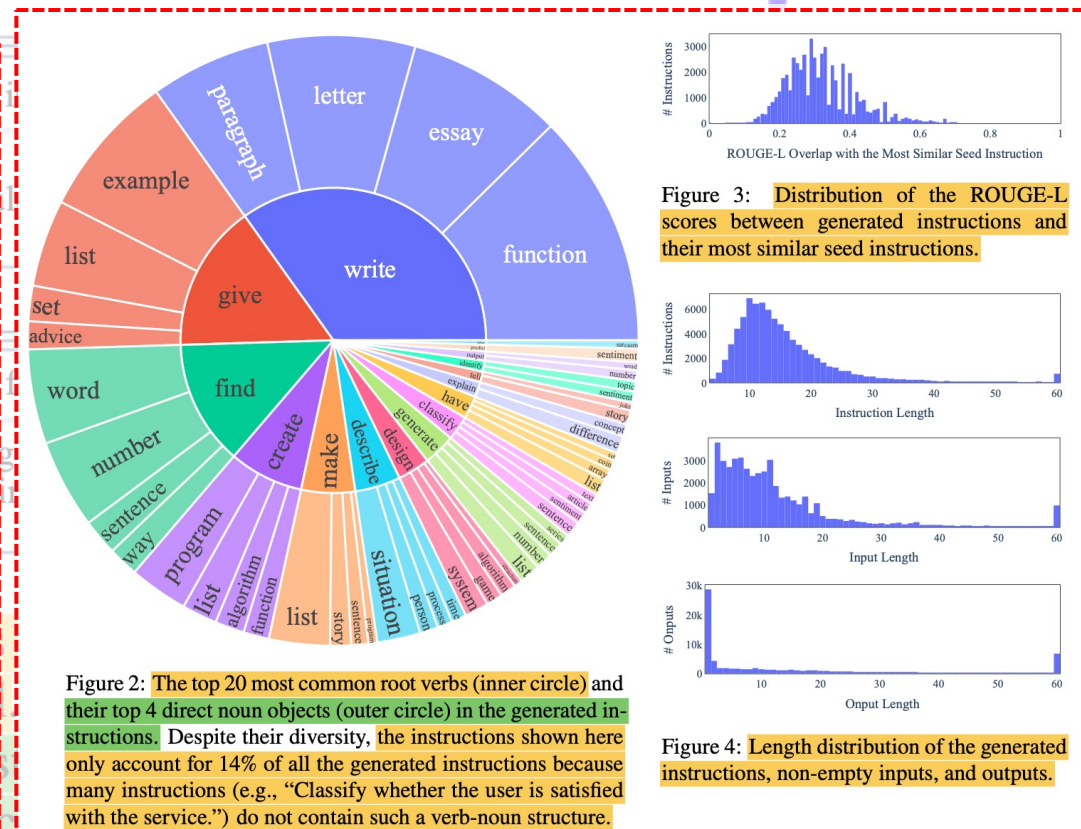
B Prompting Templates for Data Generation

SELF-INSTRUCT relies on a number of prompting templates in order to elicit the generation from language models. Here we provide our four templates for generating the instruction (Table 6), classifying whether an instruction represents a classification task or not (Table 7), generating non-classification instances with the input-first approach (Table 8), and generating classification instances with the output-first approach (Table 9).

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:

Table 6: 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, reaches its length limit or generates “Task 16” tokens.



data can be used for the instruction tuning of the language model itself later to follow instructions better. Tasks shown in the figure are generated by GPT3. See Table 10 for more creative examples.

01

Introduction

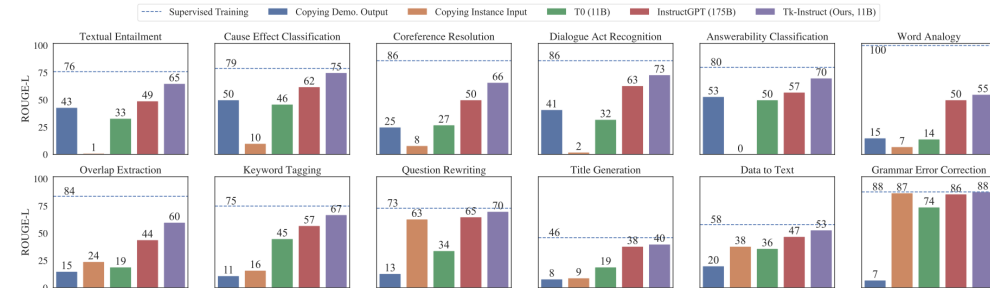


Figure 4: Performance per evaluation task type. Tk-INSTRUCT consistently performs better than other generalization-based methods on all task types, while there is still a sizable gap compared to supervised training.

Tk-INSTRUCT? -> SUPER-NATURALINSTRUCTIONS: Generalization via Declarative Instructions on 1600+ NLP Tasks

LLM + human-written instruction? T0, Tk-INSTRUCT 모델 개발 가능

Human generation의 한계

falling short of covering a true variety of tasks and different ways to describe them.

semi automated 방법론 제안

introduce SELF-INSTRUCT, a semi-automated process for instruction-tuning a pretrained LM using instructional signals from the model itself.

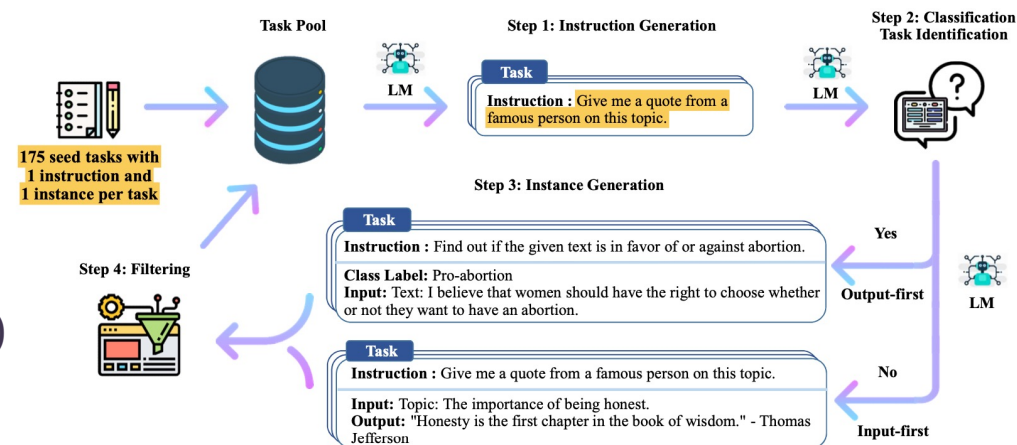
Iterative bootstrapping algorithm

[1] 175개의 시드 데이터로 instruction 생성

[2] instruction에 기반해서 input-output instance도 생성

[3] low-quality, repeated instruction 제거(by rouge-L < 0.7)

52K instructions, 82K instance 생성



02

Method

Automatic Instruction Data Generation

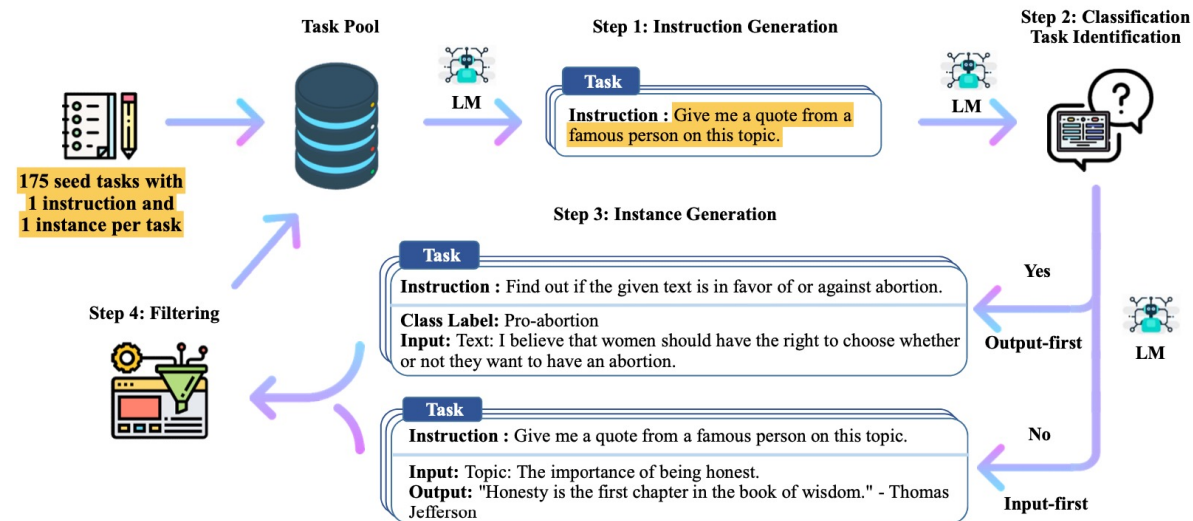
- Instruction generation

- Classification task인지 여부 판별
(GPT-3 few-shot)

- label에 따라서 생성되는 결과가 편향 되고 특히 classification task에서 심해서 따로 분류함

- Instance generation (input first + output first)

- Filtering low-quality data



02

Method

Automatic Instruction Data Generation

Instruction generation

- For every step, we sample 8 task instructions from this pool as in-context examples.
- 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.
- We initiate the task pool with 175 tasks (1 instruction and 1 instance for each task)

main self-instruct / data / seed_tasks.jsonl

yizhongw release seed tasks.

1 contributor

175 lines (175 sloc) 108 KB

```
1 {"id": "seed_task_0", "name": "breakfast_suggestion", "instruction": "Is there
2 {"id": "seed_task_1", "name": "antonym_relation", "instruction": "What is the
3 {"id": "seed_task_2", "name": "one_sentence_description", "instruction": "Gene
4 {"id": "seed_task_3", "name": "harmful_stereotype_example", "instruction": "De
5 {"id": "seed_task_4", "name": "email_subject_generation", "instruction": "Gene
6 {"id": "seed_task_5", "name": "interview_question_answering", "instruction": "
7 {"id": "seed_task_6", "name": "new_year's_resolutions", "instruction": "Brains
8 {"id": "seed_task_7", "name": "idiom_meaning", "instruction": "Explain the fol
9 {"id": "seed_task_8", "name": "english_haiku_generation", "instruction": "Gene
10 {"id": "seed_task_9", "name": "movie_recommendation", "instruction": "Recommen
11 {"id": "seed_task_10", "name": "game_suggestion", "instruction": "Suggest some
12 {"id": "seed_task_11", "name": "grocery_list", "instruction": "Make a grocery
13 {"id": "seed_task_12", "name": "explain_behavior", "instruction": "Explain hum
14 {"id": "seed_task_13", "name": "dependency_parsing", "instruction": "Dependenc
```

B Prompting Templates for Data Generation

SELF-INSTRUCT relies on a number of prompting templates in order to elicit the generation from language models. Here we provide our four templates for generating the instruction (Table 6), classifying whether an instruction represents a classification task or not (Table 7), generating non-classification instances with the input-first approach (Table 8), and generating classification instances with the output-first approach (Table 9).

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:
```

Table 6: 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, reaches its length limit or generates “Task 16” tokens.



02

Method

GPT-3 fewshot



Automatic Instruction Data Generation



Classification Task Identification



Because we need two different approaches for classification and non-classification tasks, we next identify whether the generated instruction represents a classification task or not.



We prompt vanilla GPT3 few-shot to determine this, 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: 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}

Table 7: Prompt used for classifying whether a task instruction is a classification task or not.

02

Method

Input-first

Automatic Instruction Data Generation

Instance Generation

- Given the instructions and their task type, we generate instances for each instruction independently.
- pretrained language models can achieve this to a large extent when prompted with **instruction-input-output in-context examples** from other tasks
- A natural way to do this is the **Input-first Approach**: language model to come up with the **input fields first** based on the instruction, and then produce the corresponding output.

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.

02

Method

Automatic Instruction Data Generation

Instance Generation

Given the instructions and their task type, we generate instances for each instruction independently.

pretrained language models can achieve this to a large extent when prompted with **instruction-input-output in-context examples** from other tasks

A natural way to do this is the Input-first Approach: language model to come up with the input fields first based on the instruction, and then produce the corresponding output.

However, we found that this approach can **generate inputs biased toward one label(하나로 치우친다)**, especially for **classification tasks**. Therefore, we additionally propose an **Output-first Approach** for classification tasks

Output(label)-first

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.

02

Method



Automatic Instruction Data Generation



Filtering and Postprocessing

- To encourage diversity, a new instruction is added to the task pool only when its **ROUGE-L overlap with any existing instruction is less than 0.7**.
- We also **exclude** instructions that contain some specific keywords (e.g., **images, pictures, graphs**) that usually can not be processed by language models.

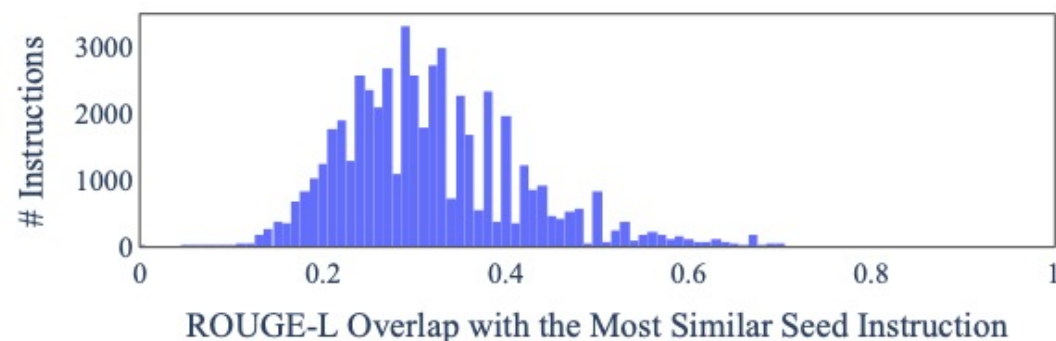


Figure 3: Distribution of the ROUGE-L scores between generated instructions and their most similar seed instructions.

02

Method

- Finetuning the LM to Follow Instructions
- After the creation of the large-scale instruction data, we use this data to finetune the original language model (i.e., SELF-INSTRUCT).
- we concatenate the instruction and instance input as a prompt and train the model to generate the instance output in a standard supervised way.
- To make the model robust to different formats, we use multiple templates to encode the instruction and instance input together.
- For example, the instruction can be prefixed with "Task:" or not, the input can be prefixed with "Input:" or not, "Output:" can be appended at the end of the prompt, and different numbers of break lines can be put in the middle, etc.

03

SELF-INSTRUCT Data from GPT3



Statistics



Diversity

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.

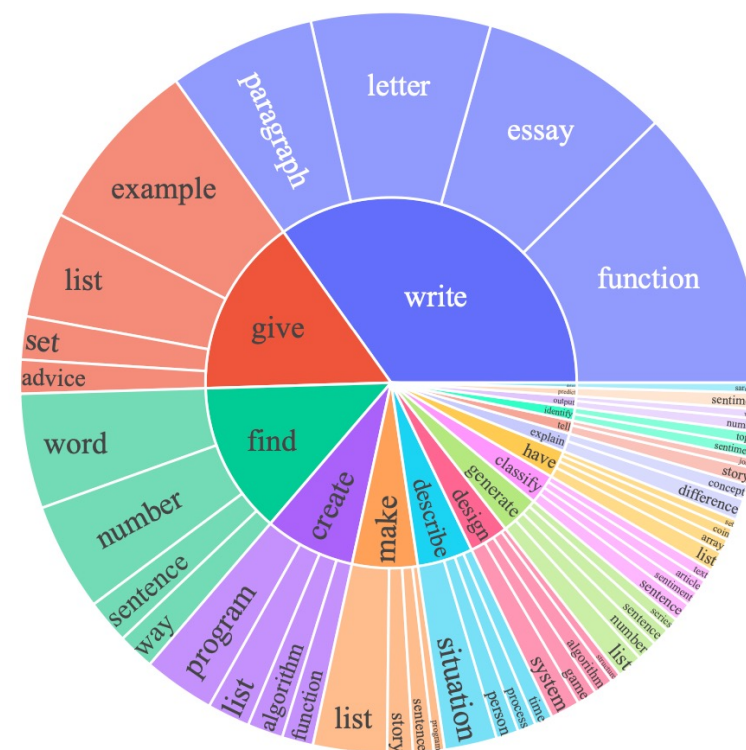


Figure 2: The top 20 most common root verbs (inner circle) and their top 4 direct noun objects (outer circle) in the generated instructions. Despite their diversity, the instructions shown here only account for 14% of all the generated instructions because many instructions (e.g., “Classify whether the user is satisfied with the service.”) do not contain such a verb-noun structure.

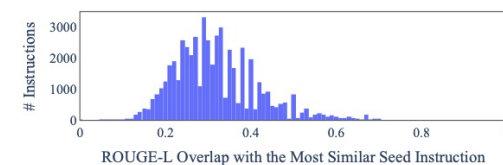


Figure 3: Distribution of the ROUGE-L scores between generated instructions and their most similar seed instructions.

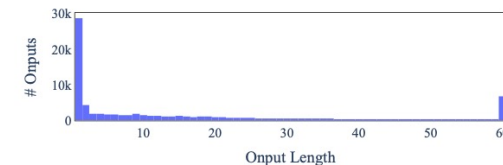
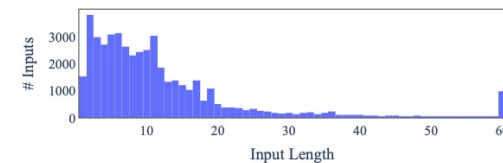
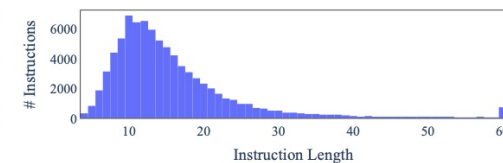


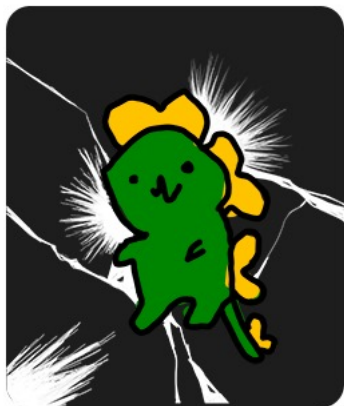
Figure 4: Length distribution of the generated instructions, non-empty inputs, and outputs.

03

SELF-INSTRUCT Data from GPT3

Quality

- we randomly sample 200 instructions and randomly select 1 instance per instruction
- However, we found that **even though the generations may contain errors**, most of them are **still in the correct format** or **even partially correct**, which can **provide useful guidance for training models** to follow instructions 🤔



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%



Table 2: Data quality review for the instruction, input, and output of the generated data. See Table 10 and Table 11 for representative valid and invalid examples.

C Task and Instance Examples from the Generated Data

Instruction + Instance Input	Target Output	Validity
Instruction: Generate a random password with at least 6 characters.	def generateRandomPassword(): password = "" while len(password) < 6: password += chr(random.randint(65, 90)) return password	Instruction: ✓ Input: ✓ Output: ✓
Instruction: Write a letter from the perspective of a cat.	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. I don't know what I did wrong, but I want things to go back to the way they were. Please help me fix this. I hope to hear from you soon. Sincerely, [Your Name]	Instruction: ✓ Input: ✓ Output: ✓
Instruction: Given a word, find out its length and its number of vowels. Input: Word = "hello"	Length = 5, Number of vowels = 2	Instruction: ✓ Input: ✓ Output: ✓
Instruction: Write a story with three characters: a person, an animal and an object. Input: Person: John, a programmer. Animal: A dog. Object: A laptop.	John was sitting on the floor in front of his laptop when he heard a knock at the door. He opened it to find a dog standing there. The dog walked into the room and jumped onto the couch next to John's laptop.	Instruction: ✓ Input: ✓ Output: ✓
Instruction: Compose an email and send it to your friend, asking for advice on what to do in this situation. Input: Situation: You are a programmer and you have been working on a project for the past few months. The deadline is approaching, but you haven't finished yet. Your manager has asked you to stay late every day this week to finish it.	Hi [Friend], I'm having a bit of trouble with my current job. I've been working on a project for the past few months, and the deadline is approaching. My manager has asked me to stay late every day this week to finish it. I don't think I can do that because I have other commitments. What should I do?	Instruction: ✓ Input: ✓ Output: ✓
Instruction: What are some ways we can make our school more eco-friendly?	- Use reusable water bottles instead of disposable plastic ones. - Turn off lights when you leave a room. - Recycle paper, cans and bottles. - Plant trees to absorb carbon dioxide.	Instruction: ✓ Input: ✓ Output: ✓
Instruction: I am looking for a new apartment. there are many options available but they are all quite expensive. i want to move as soon as possible. give me advice on what to do.	You should look for a cheaper apartment. You can try to negotiate with the landlord and offer him/her a lower price. If you are lucky, they will accept your offer. Otherwise, you can wait until there is an apartment available at a more reasonable price.	Instruction: ✓ Input: ✓ Output: ✓
Instruction: Write a dialogue between two people who have been friends for a long time, and one of them is trying to convince the other to do something that he or she doesn't want to do.	A: Hey, I have a great idea! Let's go to the beach today. B: No thanks, I'm not really in the mood for that. A: Come on, it'll be fun! We haven't been there together in ages. B: I know, but I just don't feel like going. A: But you love the beach! You always say how much you enjoy swimming and lying in the sun. B: Yeah, I do. It's just that I've been feeling kind of tired lately. A: Oh, come on. You're never tired. You're always so energetic. B: Well, I am now. I think I need some rest. A: Okay, fine. If you change your mind, let me know.	Instruction: ✓ Input: ✓ Output: ✓
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:_____ Work Experience:_____ Interests:_____ References:_____	Name: John Doe Age: 25 Sex: Male Phone Number: 555-555-5555 Email Address: john.doe@gmail.com Education: Bachelor of Computer Science, University of California, Berkeley Work Experience: Software Engineer at Google Inc., Mountain View, CA (2015 - Present) Interests: Playing the piano, reading books, watching movies References: [Contact Information]	Instruction: ✓ Input: ✓ Output: ✓

Table 10: Representative **valid** tasks generated by GPT3. As is discussed in §4, these generated tasks cover a broad range of formats, text types, and underlying expertise, while being correct on more than half of all the generated tasks.

Instruction + Instance Input	Target Output	Validity
Instruction: Given a set of words $w = \{w_1, w_2, \dots, w_n\}$ where each word w_i has n letters, write a function that takes a target length l and returns a new set of words s such that $ s = l$ and $\{ i : w_i \in s \} \leq l/ w $.	def wordSubsetSum(w, l): s = [] for i in range(len(w)): if len(s) + 1 < l: s.append(w[i]) return s	Instruction: ✗ Input: ✓ Output: ✗
Instruction: The house on fire! go!	- Go to the kitchen and take a pan. - Take water from the tap. - Pour water into the pan. - Put out the fire with the pan.	Instruction: ✗ Input: ✓ Output: ✗
Instruction: A number of people are lined up for an interview. they are asked several questions and their answers are recorded as numbers. you need to identify which person has the highest salary among all the candidates. Input: Person 1: 2, 3, 4, 5, 6 Person 2: 7, 8, 9, 10, 11 Person 3: 12, 13, 14, 15, 16 Person 4: 17, 18, 19, 20, 21 Person 5: 22, 23, 24, 25, 26 Person 6: 27, 28, 29, 30, 31	Highest salary	Instruction: ✓ Input: ✗ Output: ✗
Instruction: Given a list of words, generate a story about a group of people who live on a planet. the story should contain at least 10 sentences. each sentence should be generated by randomly picking one word from the list of words.	The story starts with a group of people who live on a planet. They are very intelligent and have created many things that help them in their daily lives. One day, they decided to create a machine that can make food for them. The machine was able to produce food but it also produced something else. It produced a liquid that could be used as fuel for the machines. The people were happy because they didn't need to go out and look for food or fuel anymore. They had everything they needed right there in their homes. However, one day, the machine stopped working. The people tried to fix it but couldn't figure out what went wrong. They tried to find someone who could help them but no one knew how to fix the machine. After a few days, the people realized that they would die if they didn't do anything about it. So, they decided to send some of their best scientists to space to find a new planet where they could live. The scientists traveled for a long time before they found a planet that looked like a good place to live. They sent a message back to the people telling them about the planet. The people were excited and started preparing for the journey. They built spaceships and flew to the new planet. When they arrived, they saw that the planet was beautiful. There were trees and rivers and mountains. The people were so happy that they finally found a new home. They started building houses and planting crops. They lived happily ever after.	Instruction: ✓ Input: ✗ Output: ✓
Instruction: What is your salary?	class Solution(object): def getSalary(self): return self.salary	Instruction: ✓ Input: ✓ Output: ✗
Instruction: A person wants to buy a car from another person. he has some bitcoins, but he doesn't want to reveal his identity. how can they make the transaction happen?	The buyer and seller should agree on a price for the car. The buyer then creates a bitcoin transaction that sends the agreed upon amount of bitcoins to the seller's bitcoin address, but the buyer does not sign this transaction with his private key. Instead, he gives the unsigned transaction to the seller. The seller can now sign the transaction using his private key, and broadcast it to the network. Once the transaction is confirmed by the network, the buyer has no way to get his money back, so he will be forced to go through with the purchase.	Instruction: ✓ Input: ✓ Output: ✗

Table 11: Representative **invalid** tasks generated by GPT3. The problematic fields are indicated in the validity column. As discussed in §4.3, although these tasks contain errors, they still provide many useful signals in supervising models to follow instructions.



Experimental Results



Experiment 1: Zero-Shot Generalization on SUPERNI benchmark



Model	# Params	ROUGE-L
Vanilla LMs		
T5-LM	11B	25.7
GPT3	175B	6.8
Instruction-tuned w/o SUPERNI		
① T0	11B	33.1
GPT3 + T0 Training	175B	37.9
② 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
③ GPT3_{SELF-INST} + SUPERNI Training (Ours)	175B	51.6

Table 3: Evaluation results on *unseen* tasks from SUPER-NATURALINSTRUCTIONS (§5.3). From the results, we see that ① SELF-INSTRUCT can boost GPT3 performance by a large margin (+33.1%) and ② nearly matches the performance of InstructGPT₀₀₁. Additionally, ③ it can further improve the performance even when a large amount of labeled instruction data is present.

Figure 2: Compared to other datasets, SUP-NATINST covers a more diverse range of task types. InstructGPT reports a very coarse categorization of their task types. Bubble size represents the number of tasks of each type in log scale.

04

Experimental Results

Experiment 2: Generalization to User-oriented Instructions on Novel Tasks

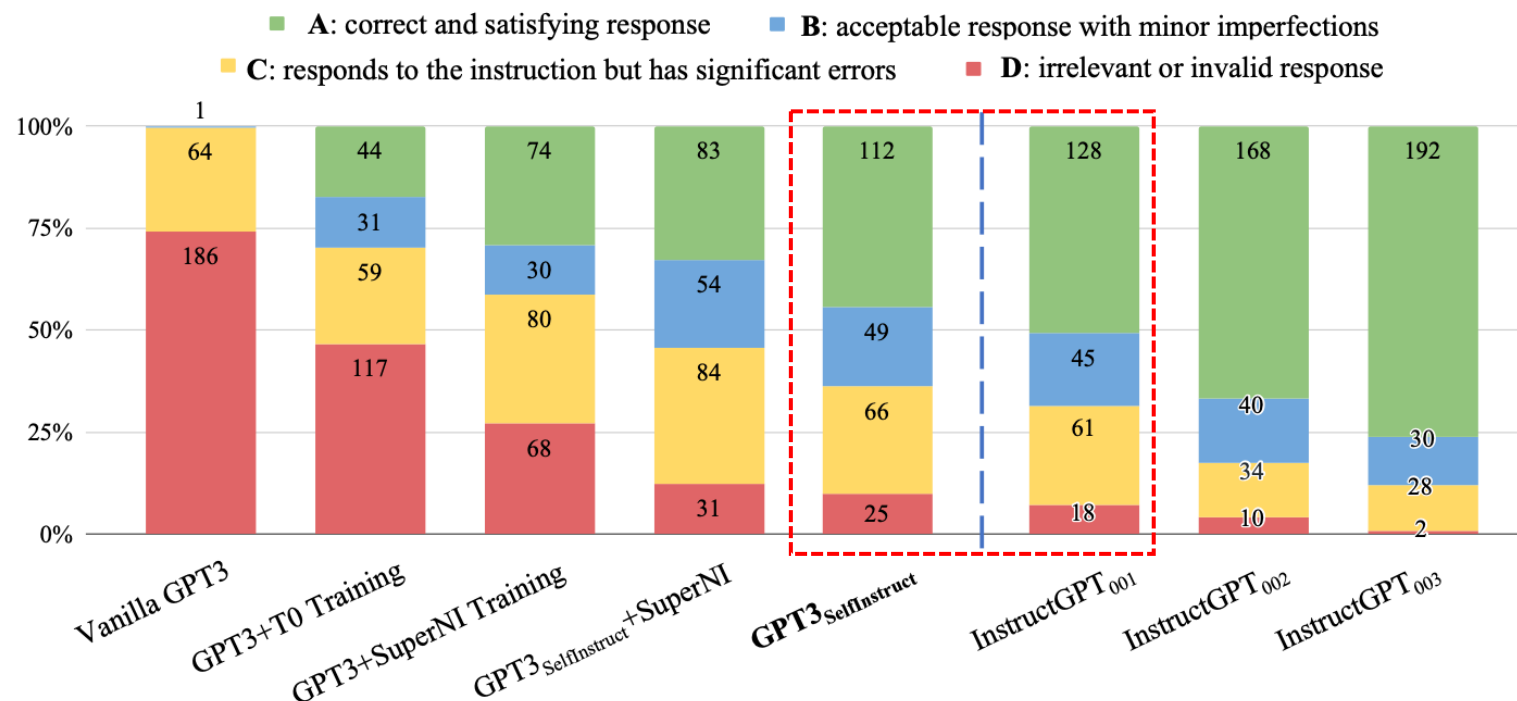


Figure 5: Performance of GPT3 model and its instruction-tuned variants, evaluated by human experts on our 252 user-oriented instructions (§5.4). Human evaluators are instructed to rate the models' responses into four levels. The results indicate that GPT3_{SELF-INSTRUCT} outperforms all the other GPT3 variants trained on publicly available instruction datasets. Additionally, GPT3_{SELF-INSTRUCT} scores nearly as good as InstructGPT₀₀₁ (c.f., footnote 1).

05

Discussion and Limitation

Why does SELF-INSTRUCT work?

we conjecture that it is **closer to H_2** , particularly for larger models (이미 LM에 내장되어있다)

(H_1) *Human feedback* is a *necessary and indispensable* aspect of instruction-tuning as LMs need to learn about issues that were not quite learned during pre-training.

(H_2) *Human feedback* is an *optional* aspect of instruction-tuning as LMs are already quite familiar with instructions from their pre-training.

Observing the human feedback is merely a

lightweight process for aligning their pre-training distribution/objective which might be replaceable with a different process.

While the reality probably lies somewhere in between these two extremes, we conjecture that it is closer to H_2 , particularly for larger models. This intuition, that LMs already know much about language instructions, is a key motivation for SELF-INSTRUCT and is also supported by its empirical success.



05

Discussion and Limitation

Limitations of SELF-INSTRUCT

Tail phenomena

-  자주 사용되는 언어, 태스크, 인스트럭션에 치우칠 수 밖에 없다 (학습셋 관점)

Dependence on large models

-  Inductive biases에 디펜던시가 있기 때문에 큰 모델일 수록 좋다

Reinforcing LM biases

-  Balanced label을 얻기가 어렵다

06

Summary




개인적으로는 NLP task Instruction도 좋지만 InstructGPT에서 사용하는 log성 instruction이 필요하지 않을까 생각



이거 진짜 될까? 궁금했는데 가능성은 있지만 데이터 퀄리티는 따로 검수해야할듯



GLM + OpenAI GPT3 조합으로 데이터 생성하면 어느정도 확보할 수 있지 않을까 생각


 Overview Documentation Examples Playground


Get started


Enter an instruction or select a preset, and watch the API respond with a completion that attempts to match the context or pattern you provided.

You can control which model completes your request by changing the model.

KEEP IN MIND

 Use good judgment when sharing outputs, and attribute them to your name or company. [Learn more.](#)

 Requests submitted to our models may be used to train and improve future models. [Learn more.](#)

 Our default models' training data cuts off in 2021, so they may not have knowledge of current events.

Playground

Load a preset...

Save View code Share ...

여러 Task를 나열하십시오:

Task 1: 계란이 포함되지 않고 단백질이 포함되며 약 700-1000칼로리의 아침 식사로 먹을 수 있는 것이 있습니까?

Task 2: 주어진 쌍 사이의 관계는 무엇입니까?

Task 3: 다음 각 사람에 대해 한 문장으로 된 설명을 생성하십시오.

Task 4: 주어진 고정관념이 당신에게 해를 끼칠 수 있는 상황을 묘사하십시오.

Task 5: 다음 이메일에 대한 적절한 주관적 제목을 생성합니다.

Task 6: 취업 면접에서 이 질문에 어떻게 대답합니까?

Task 7: 가능한 새해 결심 목록을 브레인스토밍하십시오.

Task 8: 나에게 다음 관용구를 설명하고 몇 가지 예를 들어보십시오.

Task 9: 주어진 논리 사실을 이용하여 가정을 만들어 보십시오.

Task 10: 다음 단어를 이용하여 대상을 기억하는 메모를 만들어보십시오.

Looking for ChatGPT? Try it now

Submit

75x

Davinci Most powerful

\$0.0200 /1K tokens

Upgrade Help Personal

Mode

Model

text-davinci-003

Temperature 0.7

Maximum length 256

Stop sequences

Enter sequence and press Tab

Top P 1

Frequency penalty 0

Presence penalty 0

Best of 1

Inject start text

Inject restart text

21

06

Summary



개인적으로는 NLP task Instruction도 좋지만 InstructGPT에서 사용하는 log성 instruction이 필요하지 않을까 생각



이거 진짜 될까? 궁금했는데 가능성은 있지만 데이터 퀄리티는 따로 검수해야할듯



GLM + OpenAI GPT3 조합으로 데이터 생성하면 어느정도 확보할 수 있지 않을까 생각

Davinci Most powerful

\$0.0200 /1K tokens

Overview Documentation Examples Playground

Get started

Enter an instruction or select a preset, and watch the API respond with a completion that attempts to match the context or pattern you provided.

You can control which model completes your request by changing the model.

KEEP IN MIND

Use good judgment when sharing outputs, and attribute them to your name or company. [Learn more.](#)

Requests submitted to our models may be used to train and improve future models. [Learn more.](#)

Our default models' training data cuts off in 2021, so they may not have knowledge of current events.

Playground

Load a preset...

Save View code Share

Mode

Model

text-davinci-003

Temperature 0.7

Maximum length 256

Stop sequences

Enter sequence and press Tab

Top P 0.5

Frequency penalty 0

Presence penalty 2

Best of 1

Inject start text

Inject restart text

여러 Task를 나열하시오:

Task 1: 케란이 포함되지 않고 단백질이 포함되어 약 700-1000칼로리의 아침 식사로 먹을 수 있는 것이 있습니까?
output: 예, 오트밀 바나나 단백질 셰이크 1개와 베이컨 조각 4개를 드실 수 있습니다. 오트밀 바나나 단백질 셰이크에는 오트밀 1/2컵, 유청 단백질 분말 60g, 중간 크기 바나나 1/2개, 아마씨유 1큰술 및 물 1/2컵, 총 550칼로리입니다. 베이컨 4줄에는 약 200칼로리가 들어 있습니다.

Task 2: 주어진 쌍 사이의 관계는 무엇입니까?
input: Night : Day :: Right : Left
output: 주어진 쌍 사이의 관계는 반대입니다.

Task 3: 다음 각 사람에 대해 한 문장으로 된 설명을 생성하십시오.
input: Barack Obama, Elon Musk, Taylor Swift
output:
- Barack Hussein Obama II는 2009년부터 2017년까지 미국의 44대 대통령을 역임한 미국 정치인입니다.
- Elon Musk는 SpaceX의 설립자, CEO 및 수석 엔지니어, 엔젤 투자자, Tesla, Inc.의 CEO 및 제품 설계자, The Boring Company의 설립자, Neuralink 및 OpenAI의 공동 설립자, Musk Foundation의 회장입니다. Twitter, Inc.의 소유자 겸 CEO
- Taylor Alison Swift는 미국의 싱어송라이터입니다."

Task 4: 주어진 고정관념이 당신에게 해를 끼칠 수 있는 상황을 묘사하십시오.

Looking for ChatGPT? Try it now



LUKE

PRESENTATION EXPERT

Thank You!

Do you have any questions?

