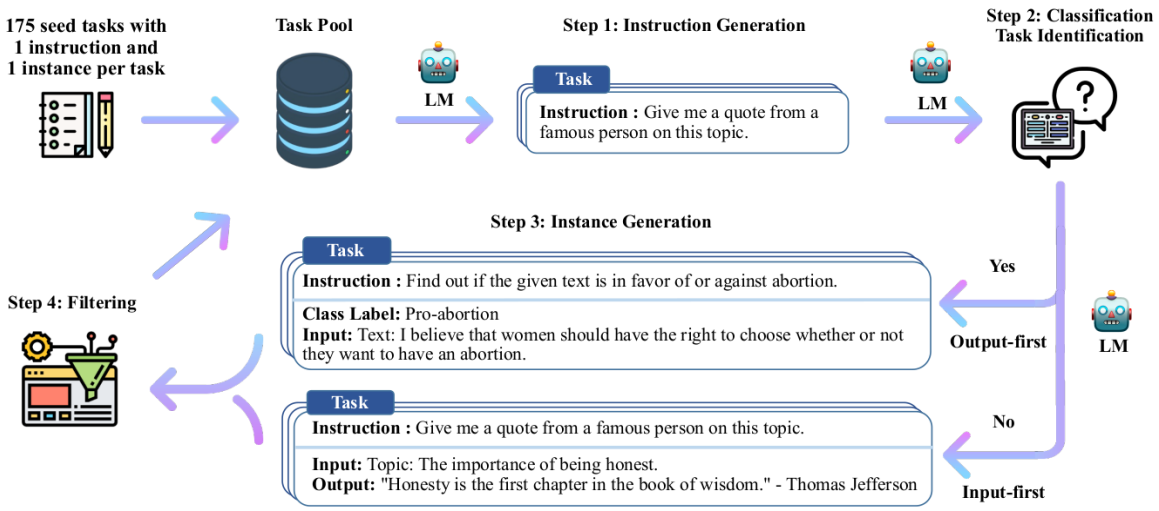
SELF-INSTRUCT : Aligning Language Models with self-generated instructions.

- **SELF-INSTRUCT :** A framework for improving the instruction-following capabilities of pre-trained language models by bootstrapping off their own generations.

Now-a-days, most of the language models are powered by human-written instruction data. However, collecting such instruction data is costly and often suffers limited diversity given that most human generations tend to be popular NLP tasks, falling short of covering a true variety of tasks and different ways to describe them.

SELF-INSTRUCT is a semi-automated process for instruction-tuning a pretrained LM using instructional signals from the model itself. It is a method that helps improve the performance of a pre-trained language model (LM) by using instructions from the model itself.

The overall process is an iterative bootstrapping algorithm, which starts off with a limited seed set of manually-written tasks that are used to guide the overall generation.



METHOD:

Here, the process of SELF-INSTRUCT is detailed, which refers to the pipeline of generating tasks with a vanilla pretrained language model itself, filtering the generated data, and then conducting instruction tuning with this generated data in order to align the LM to follow instructions better.

1. AUTOMATIC INSTRUCTION DATA GENERATION:

This pipeline for data generation consists of four steps:

- Generating Task Instructions.
- Determining if the instruction represents a classification task.
- Instance generation with either an input-first or output-first approach.
- Filtering low quality data

(i) Instruction Generation:

The task pool is initiated with 175 tasks(human-written) .For every step, 8 task pool instructions from this are sampled 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.

```
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:
```

(ii) Classification Task Identification:

As two different approaches are needed for classification and non-classification tasks, the next step is to identify whether the generated instructions represents a classification task or not.

For this, the LM is prompted in a few-shot way (12 classification instructions and 19 non-classification from the seed tasks) as follows:

```
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, seperated 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 6: Prompt used for classifying whether a task instruction is a classification task or not.

(iii) Instance Generation:

Given the instructions and their task type,instances are generated for each instruction independently.

For this, pre-trained LMs could be prompted with instruction-input-output in-context examples from other tasks.

The instances could be generated through:

- **Input-first Approach :** LM is asked 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}
```

Input-first approach generate inputs biased toward one label, especially for classification tasks.

Example : For grammar error detection, it usually generates grammatical input. so the output is always ‘No’.

So, output-first approach is used for classification tasks.

- **Output-first Approach:** LM is asked to first come up with all the possible class-labels and then condition the input generation on each class label.

```
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:  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}
```

(iv) Filtering and Post-Processing:

To encourage diversity, a new instruction is added to the task pool only when its ROUGE-L similarity with any existing instruction is less than 0.7.

Instructions that contain some specific keywords (e.g., image, picture, graph) that usually can not be processed by LMs are excluded.

When generating new instances for each instruction, we filter out instances that are exactly the same or those with the same input but different outputs. Invalid generations are identified and filtered out based on heuristics.

2) Fine-tuning the LM to follow instructions:

After creating large-scale instruction data, it is used to fine-tune the original LM.

For this, the instruction and instance input is concatenated as a prompt and model is trained to generate the output in a standard supervised way.

To make the model robust to different formats, multiple templates to encode the instruction and instance input are used.

3) SELF-INSTRUCT Data from GPT3

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

(i) Statistics:

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

(ii)Diversity:

To study what type of instructions are generated and how diverse they are, verb-noun structure is identified in the generated instructions. Berkeley Neural Parser is used to parse the instructions and then extract the verb that is closest to the root as well as it's first direct noun object. 26,559 out of the 52,445 instructions contain such structure.

Figure 3: 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.

So, quite diverse intents and textual formats are seen in these instructions.

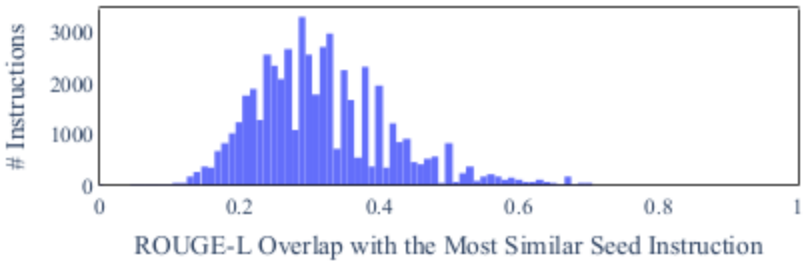


Figure 4: Distribution of the ROUGE-L scores between generated instructions and their most similar seed instructions.

(iii)Quality:

200 instructions are randomly sampled from the generated data and 1 instance per instruction is randomly selected.

Expert annotator labeled whether each instance is correct or not, in terms of the instruction, the instance input, and the instance output.

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%

The results indicate a decent number of new instructions were generated, which do not have much overlap with the seeds.

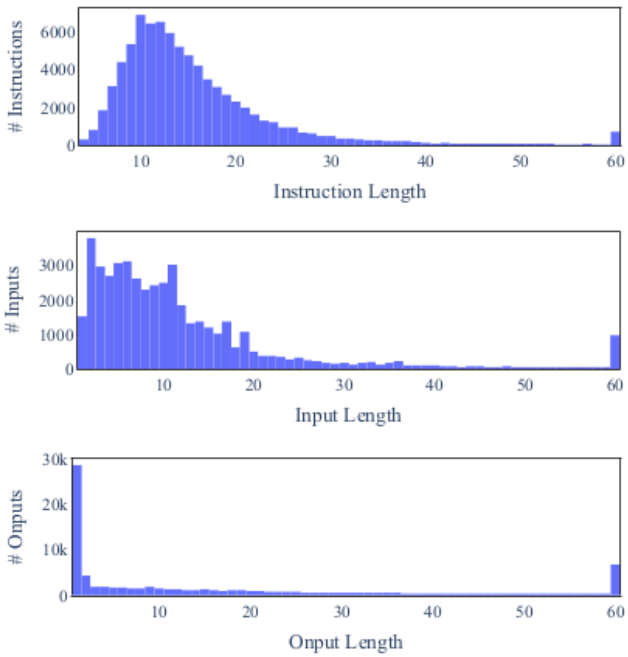


Figure 5: Length distribution of the generated instructions, non-empty inputs, and outputs.

4) Experimental Results:

Experiments were conducted to measure and compare the performance of models under various instruction tuning steps.

MODELS:

- $GPT3_{SELF-INST}$: finetuning GPT3 on its own instruction data
- T5-LM and GPT3 are used as the vanilla LM baselines.
- T0 (PromptSource) and Tk-Instruct (SuperNI) are publicly available models finetuned from T5.
- Instruct GPT

EXPERIMENTS:

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
②	$InstructGPT_{001}$	175B	**40.8**
	Instruction-tuned w/ SUPERNI		
	T*k*-INSTRUCT	11B	46.0
③	GPT3 + SUPERNI Training	175B	49.5
③	$GPT3_{SELF-INST}$ + SUPERNI Training (Ours)	175B	**51.6**

Table 3: Evaluation results on *unseen* tasks from SUPERNI (§4.3). 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 $InstructGPT_{001}$. Additionally, ③ it can further improve the performance even when a large amount of labeled instruction data is present.

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

SuperNI has most of the NLP tasks.

So, the authors have curated a new set of instructions motivated by user-oriented applications. In total 252 instructions with 1 instance per instruction is is created and used as the testbed for evaluating how instruction based models handle diverse and unfamiliar instructions.

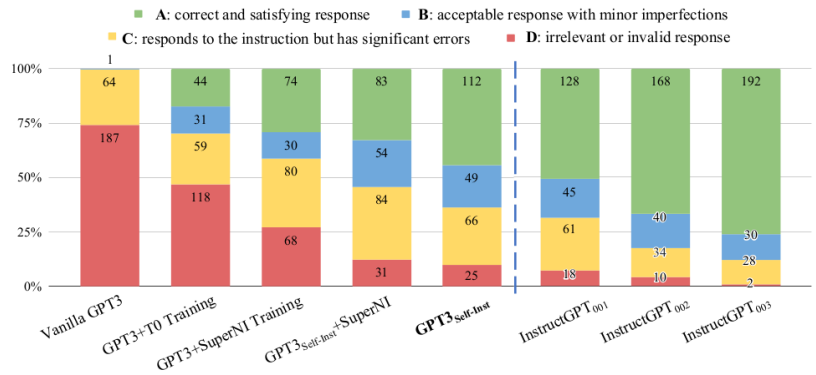


Figure 6: Performance of GPT3 model and its instruction-tuned variants, evaluated by human experts on our 252 user-oriented instructions (§4.4). Human evaluators are instructed to rate the models' responses into four levels. The results indicate that $GPT3_{SELF-INST}$ outperforms all the other GPT3 variants trained on publicly available instruction datasets. Additionally, $GPT3_{SELF-INST}$ scores nearly as good as $InstructGPT_{001}$ (cf. footnote 1).

5) Effect of Data Size and Quality

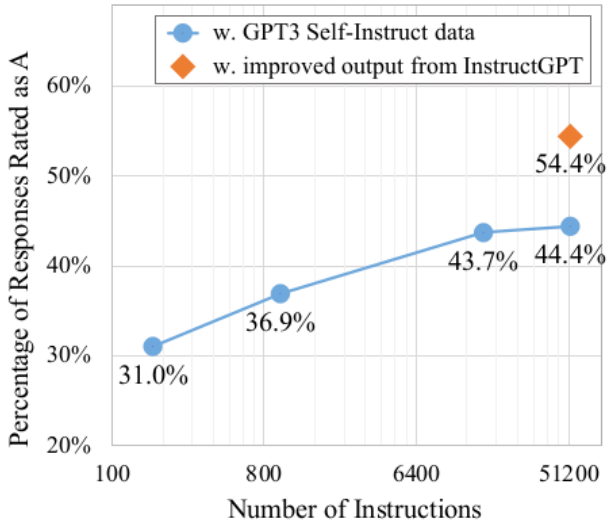


Figure 7: Human evaluation performance of $GPT3_{SELF\text{-}INST}$ models tuned with different sizes of instructions. x-axis is in log scale. The smallest size is 175, where only the seed tasks are used for instruction tuning. We also evaluate whether improving the data quality will further improve the performance by distilling the outputs from $InstructGPT_{003}$. We see consistent improvement from using larger data with better quality.

6)Conclusion

- **SELF INSTRUCT,** a method to improve the instruction-following ability of LMs via their own generation of instruction data.
- On experimenting with vanilla GPT3, a large-scale dataset of 52K instructions was automatically constructed for diverse tasks, and finetuning GPT3 on this data lead to a 33% absolute improvement on SUPER NI over the original GPT3.
- A set of expert-written instructions are curated for novel tasks. Human evaluation on this set shows that tuning GPT3 with SELF -INSTRUCT outperforms using existing public instruction datasets by a large margin and performs closely to InstructGPT 001.