Instruction Tuning of Large Language Models

Yizhong Wang
@ JHU CS 601.471/671 Class
ChatGPT/GPT4 are real generalists

Official examples

"Explain quantum computing in simple terms"

"Got any creative ideas for a 10 year old’s birthday?"

"How do I make an HTTP request in Javascript?"

Passing exams

Writing a real website for me!
How did models acquire the vast capabilities?

* Brown et al., 2020. GPT3, Ouyang et al., 2022. InstructGPT
How did models acquire the vast capabilities?

Pretraining
(GPT3*: 499 Billion tokens)

Fine-tuning
(InstructGPT*: 44K examples)

* Brown et al., 2020. GPT3, Ouyang et al., 2022. InstructGPT
Outline of this talk

● Fundamentals: cross-task generalization in NLP tasks
  ○ Super-NaturalInstructions: Generalization via Declarative Instructions on 1600+ NLP Tasks

● Replicating ChatGPT:
  ○ Self-Instruct: Aligning Language Model with Self-Generated Instructions

● Discussion of recent trends and future directions
NLP before 2018: building task-specific models

- **Sentiment Analysis**
  - “My experience has been fantastic!”
  - “Positive”

- **Question Answering**
  - “Where is World Cup 2022 playing?”
  - “Qatar.”

- **Machine Translation**
  - “AI is changing the daily lives.”
  - “人工智能正在重塑日常生活。”

Instance-level generalization within one task
Classical multi-task learning

- **Sentiment Analysis**
  - “My experience has been fantastic!”
- **Question Answering**
  - “Where is World Cup 2022 playing?”
- **Machine Translation**
  - “AI is changing the daily lives.”
  - “Positive”
  - “Qatar.”
  - “人工智能正在重塑日常生活。”

*McCann et al., 2019 DecaNLP*
Classical MTL cannot generalize to unseen tasks

Model cannot generalize to the “unseen” task because it doesn’t know what the task is.

“That’s awesome! I really love being out here with my son. Do you think you could spare some food?”
Generalization to unseen tasks via instructions

*Mishra et al., 2022 Natural Instructions V1*
Benchmarking massive MTL and cross-tasks generalization

- Super-NaturalInstructions collection: 1616 tasks in 76 broad categories
Expert-written instructions for all tasks

Definition is declarative, and should be sufficient to define the task to average human readers.

Example definition for SQuAD QA task:

This task is about writing a correct answer for the reading comprehension task. Based on the information provided in a given passage, you should identify the shortest continuous text span from the passage that serves as an answer to the given question. Avoid answers that are incorrect or provides incomplete justification for the question.
Strict train/test split for cross-task generalization

Train: 64 categories, 757 tasks
- Sentiment Analysis
- Question Answering
- Question Generation
- Dialogue Generation
- Summarization
- Grammar Error Detection
- Sentence Composition

Test: 12 categories, 119 tasks
- Textual Entailment
- Cause Effect Clf.
- Coreference
- Dialogue Act Recognition
- Answerability Clf.
- Word Analogy
- Overlap Extraction
- Keyword Tagging
- Question Rewriting
- Title Generation
- Data to Text
- Grammar Error Correction
Instruction tuning significantly improves LLMs
What are the most important factors?

- Diverse tasks, rather than more data in a single task
What are the most important factors?

- Bigger pretrained language models
What are the most important factors?

- Good instructions or in-context examples
Takeaways

● Cross-task generalization via instructions is plausible.
● Super-NaturalInstructions provides a rich playground for such study.
● For instruction tuning:
  ○ Task/Instruction diversity is important!
  ○ Larger models bring in consistent improvement - not converged yet.
  ○ Large number of training instances could lead to overfitting to the training task.
- Ok, cool. Can I try your model for my creative needs?
- Sorry, you might get disappointed.
Other models trained on existing NLP datasets

T0 (3B)

- Hosted inference API

Text2Text Generation

Write a sentence with the following words. Your output should contain all the words. Input: Apple, cash, tech.

```
Compute  Enter
```

Computation time on cpu: 2.320 s

cash in the bank ... a tech company

FLAN-T5 (3B)

- Hosted inference API

Text2Text Generation

Write a sentence with the following words. Your output should contain all the words. Input: Apple, cash, tech.

```
Compute  Enter
```

Computation time on cpu: 1.828 s

cash for tech and cash for apple
InstructGPT is great!

Write a sentence with the following words. Your output should contain all the words.
Input: Apple, cash, tech.

Apple is a tech company that deals in cash.
Data is OpenAI’s secret weapon

GPT-3

Labeler-written Prompts & Instances for training a Instruction-following model (11295)

Contractors

API Users

Contractors

Contractors

SFT

Customer Prompts

SFT

Labeler-written Desired Output

Labeler Preference over Model’s Generations

InstructGPT

A lot of creative queries from users and human labeling!
Can we construct a similar instruction dataset by crowdsourcing?

We did a pilot study but found that:

● Writing diverse instructions requires creativity.
● Writing instances for different instructions requires broad expertise.
● **Impractical** for crowd workers.
Can LLMs generate instruction data?

Pretraining
(GPT3*: 499 Billion tokens)

LLMs should know a lot of tasks!
LLMs can be prompted to generate instructions

Come up with a series of tasks:

1. I am planning a 7-day trip to Seattle. Can you make a detailed plan for me?
2. Is there anything I can eat for a breakfast that doesn’t include eggs, yet includes protein, and has roughly 700-1000 calories?
3. Translate this sentence into any Arabic dialect and say what dialect it is: “The beautiful ones are not yet born
4. Given a set of numbers, find all possible subsets that sum to a given number.
5. Give me a phrase that I can use to express I am very happy.
6. Create a birthday planning checklist.
7. What is the relation between the given pairs?
8. Tell if the sentence contains toxic language.
9. Create a list of 10 words that have the letter “k” in them.
10. Create a list of all countries and their capitals from Africa.
11. I am looking for a job, but it’s difficult to find one. Can you help me?
12. Find the longest word in this sentence: The quick brown fox jumps over the lazy dog.
13. Translate these numbers into English (word format): 3168987646561523
14. What is the next number in the series: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096...?
15. Write a program that will tell me if a given string contains anagrams.
LLMs can be prompted to generate instructions

7. What is the relation between the given pairs?
8. Tell if the sentence contains toxic language.
9. Create a list of 10 words that have the letter “k” in them.
10. Create a list of all countries and their capitals from Africa.
11. I am looking for a job, but it's difficult to find one. Can you help me?
12. Find the longest word in this sentence: The quick brown fox jumps over the lazy dog.
13. Translate these numbers into English (word format): 3168987645561523
14. What is the next number in the series: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, 4096...?
15. Write a program that will tell me if a given string contains anagrams.
LM can be prompted to generate instances

Come up with an example for each of the following task. Each example must have one output field. If the task requires input, it should be generated before the output.

Task 1: Make a list of things to do in the given city.
Input: ...
Output: ...

Task 2: Converting 85 F to Celsius.
Output: ...

Task 3: Extract all the country names in the paragraph, list them separated by comma.
Paragraph: ...
Output: ...

Task 4: Suggest a better and more professional rephrasing of the following sentence.
Sentence: ...
Output: ...

Task 5: Read the following paragraph and answer a math question about the paragraph. You need to write out the calculation for getting the final answer.
Paragraph: ...
Question: ...
Output: ...

Task 6: Solving the equation and find the value of X.
Equation: ...
Output: ...

Task 7: Write a knock knock joke about bananas.
Output: ...

Task 8: Tell me whether the given sentence is passive or not.
Sentence: The dog was bitten by the cat.
Output: Passive, because the subject of the sentence is being acted upon (the dog).
LM can be prompted to generate instances

Task 6: Solving the equation and find the value of X.
Equation: ...
Output: ...

Task 7: Write a knock knock joke about bananas.
Output: ....

Task 8: Tell me whether the given sentence is passive or not.
Sentence: The dog was bitten by the cat.
Output: Passive, because the subject of the sentence is being acted upon (the dog).
Instruction data generation pipeline

Step 1: Instruction Generation

Task

Instruction: Give me a quote from a famous person on this topic.

Step 2: Classification Task Identification

LM

Step 3: Instance Generation

Task

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

Class Label: Pro-abortion
Input: Text: I believe that women should have the right to choose whether or not they want to have an abortion.

Task

Instruction: Give me a quote from a famous person on this topic.

Input: Topic: The importance of being honest.
Output: "Honesty is the first chapter in the book of wisdom." - Thomas Jefferson
Generating 52K instructions with GPT3

- We applied self-instruct to GPT3 (`davinci` engine).
- We generated 52K instructions and 82K instances.
- ROUGE-L overlap between any pair of instructions < 0.7.
Tasks generated by GPT3

**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: 555-555-5555
...

**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:** I am looking for a new apartment. There are many options but they are all quite expensive. I want to move as soon as possible. Give me advice on what to do.
**Input:** Null
**Output:** 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 ...

**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 ...
Data quality review

- 200 random instructions are sampled for quality check

<table>
<thead>
<tr>
<th>Review Question</th>
<th>Yes %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the instruction describe a valid task?</td>
<td>92%</td>
</tr>
<tr>
<td>Is the input appropriate for the instruction?</td>
<td>79%</td>
</tr>
<tr>
<td>Is the output a correct and acceptable response to the instruction and input?</td>
<td>58%</td>
</tr>
<tr>
<td>All fields are valid.</td>
<td>54%</td>
</tr>
</tbody>
</table>
Self-instruct: finetuning GPT3 with the data generated by itself

- We finetuned GPT3 with this synthetic instruction data via their API*.
- Two epochs.
- $338 for finetuning the “davinci” engine on the 82K instances.

* still unclear about the precise details of how the API work (e.g., which parameters are updated, or which version of the GPT3 model is used).
1. Self-Instruct boosts GPT3 by 33.1%.
2. Nearly matches the perf. of InstructGPT\textsubscript{001}.
3. Complementary improvement to the existing human-labeled training set.
Expert evaluation on 252 user-oriented instructions

- **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

Here is the breakdown of the evaluation:

- **Vanilla GPT3**:
  - A: 64
  - B: 44
  - C: 187
  - D: 1

- **GPT3+T0 Training**:
  - A: 74
  - B: 31
  - C: 59
  - D: 1

- **GPT3+SuperNI Training**:
  - A: 83
  - B: 54
  - C: 80
  - D: 1

- **GPT3_Self-Inst**:
  - A: 112
  - B: 49
  - C: 31
  - D: 25

- **InstructGPT_001**:
  - A: 128
  - B: 45
  - C: 61
  - D: 18

- **InstructGPT_002**:
  - A: 168
  - B: 34
  - C: 10
  - D: 40

- **InstructGPT_003**:
  - A: 192
  - B: 30
  - C: 28
  - D: 2

Each bar represents the number of responses falling into each category for different models and training conditions.
Effect of data size and data quality (using human eval)

- More instructions -> better perf.
- Slower gain after the size of 16K
Effect of data size and data quality (using human eval)

- Improving the output quality (using InstructGPT-003) can significantly boost the performance.
Takeaways

● You don’t need a huge amount of labeled data to get good initial instruction following ability.

● LLMs themselves know many tasks/skills.

● One aspect of instruction tuning is to elicit these skills.
  ○ i.e., alignment.
  ○ Self-instruct is an extreme setup.

● Self-instruct is applicable to a broad range of LLMs, both vanilla and instruction-tuned.

● More broadly, humans & AI should collaborate in building datasets.
Recent trend: more variants of self-instruct

- The major improvements
  - Distilling from GPT3.5/ChatGPT/GPT4
  - Simplified pipeline - no special care for classification tasks
  - Better seed tasks (e.g., using ShareGPT)
  - Dialogue format
  - Finetuning LLaMa

Alpaca (Stanford)  Vicuna (UCB)  Baize (UCSD)  Koala (UCB)  Instruct-tuning with GPT4 (MSR)
Licensing concern about using OpenAI output?

**OpenAI Terms of Use**

Restrictions. You may not ... (iii) use output from the Services to develop models that compete with OpenAI...

- Ok, we can generate data using LLaMa “self”-instruct.

---

What is the difference between a linked list and an array?
Output: Linked list stores elements sequentially using pointers whereas arrays store elements contiguously in memory.

---

How would you describe yourself to someone who doesn’t know you at all?
Output: Hello, nice to meet you. My name is [your_name]. I work as a software engineer at [company_name]. I like playing basketball during my free time. I also enjoy reading books and watching movies.

---

What is the best way to get from point a to b?
Example 1
Point A: 100 Main Street, New York City, NY
Point B: 100 Wall St., New York City, NY
Output: Take subway line 4 towards Brooklyn Bridge Station, then walk southwest toward Pearl St. Turn right onto Pearl St. Then turn left onto Wall St. Your destination will be on the right. It takes around 1 hour.
Example 2
Point A: 100 Main Street, San Francisco, CA
Point B: 100 Wall St., Los Angeles, CA
Output: Drive northwest on Market St toward Montgomery St. Continue straight onto Spear St. Keep right at the fork, follow signs for US-101 N/Golden Gate Bridg and merge onto Octavia Blvd. Use the middle lane to take the ramp onto US-101 N. Follow US-101 S and I-5 S to W 7th St in Los Angeles. Exit from I-10 E. Merge onto I-5 S. Take exit 1A to merge onto CA-110 S toward Harbor Fwy. Take exit 22B to merge onto W 7th St. Destination will be on the right. It takes around 6 hours.
Licensing concern about using OpenAI output?

**OpenAI Terms of Use**
Restrictions. You may not … (iii) use output from the Services to develop models that compete with OpenAI...

- Ok, we can generate data using LLaMa “self”-instruct.
- More human-labeled data will come out soon (e.g., Free Dolly)
Promising directions

● Understanding the fundamentals of instruction tuning.
● Combining all efforts and building an open-sourced general-purpose model.
  ○ A lot of data will come.
  ○ LLaMa is a good base model, and more base models will come.
  ○ The open-source community will catch up soon.
● Algorithms for further improving the models (human feedback or self-improve)
● Specialization/personalization of general-purpose models.
Promising directions (cont.)

Pretraining

Extensive data will be generated here!

Fine-tuning (sft, rlhf, etc.)
Thanks!

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

@wyz1e32
@yizhongwyz
yizhongw@cs.washington.edu