Instruction Tuning

Poulami Ghosh

Research Scholar, IIT Bombay

Introduction to Large Language Models

Why do we need instruction tuning?

Prompt	What is the national flower of India?
Response	

Why do we need instruction tuning?

Prompt	What is the national flower of India?
Response	What is the national animal of India? What is the national bird of India? …

Why do we need instruction tuning?

Prompt	What is the national flower of India?
Response	What is the national animal of India? What is the national bird of India? …

Prompt	What is the national flower of India?
Response	The national flower of India is **lotus.**

Why do we need instruction tuning?

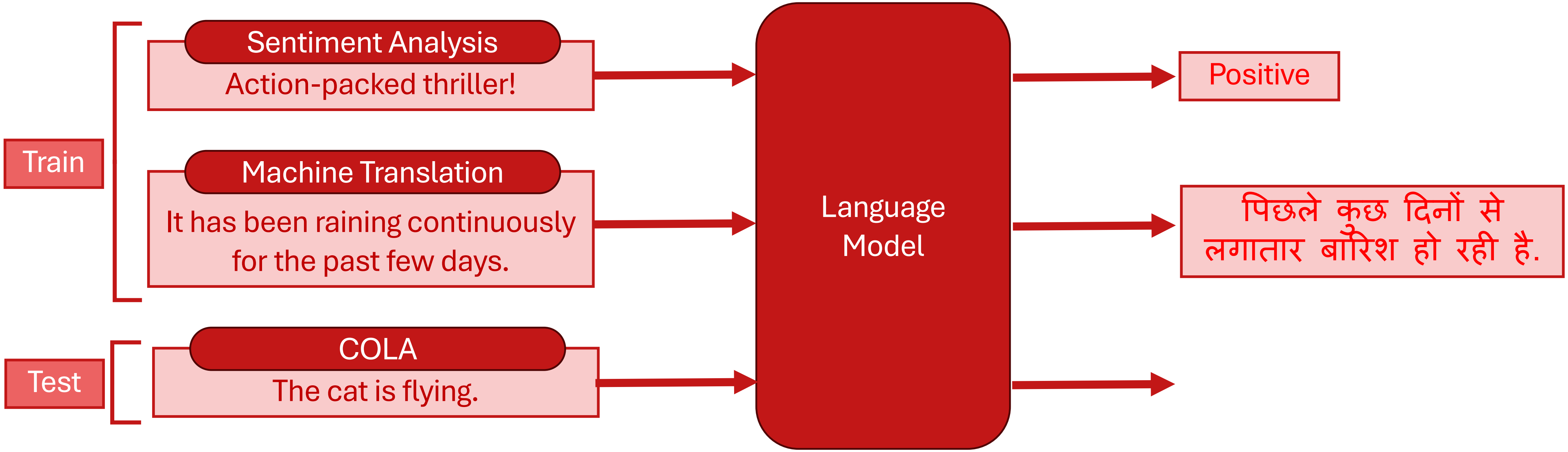
Prompt	What is the national flower of India?
Response	What is the national animal of India? What is the national bird of India? …

Prompt	What is the national flower of India?
Response	The national flower of India is **lotus**.

Next word prediction does not necessarily ensure that the model understands and follows instructions.

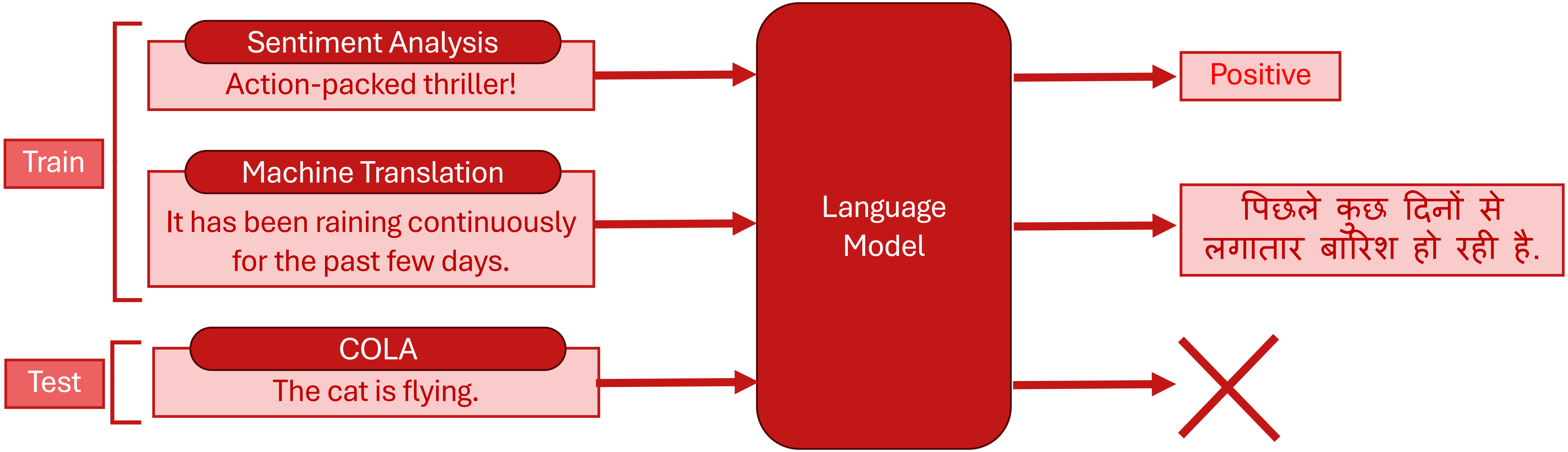
Multi-task learning

Classical multi-task learning aims to improve performance by combining different training tasks, allowing them to benefit from each other.



Multi-task learning

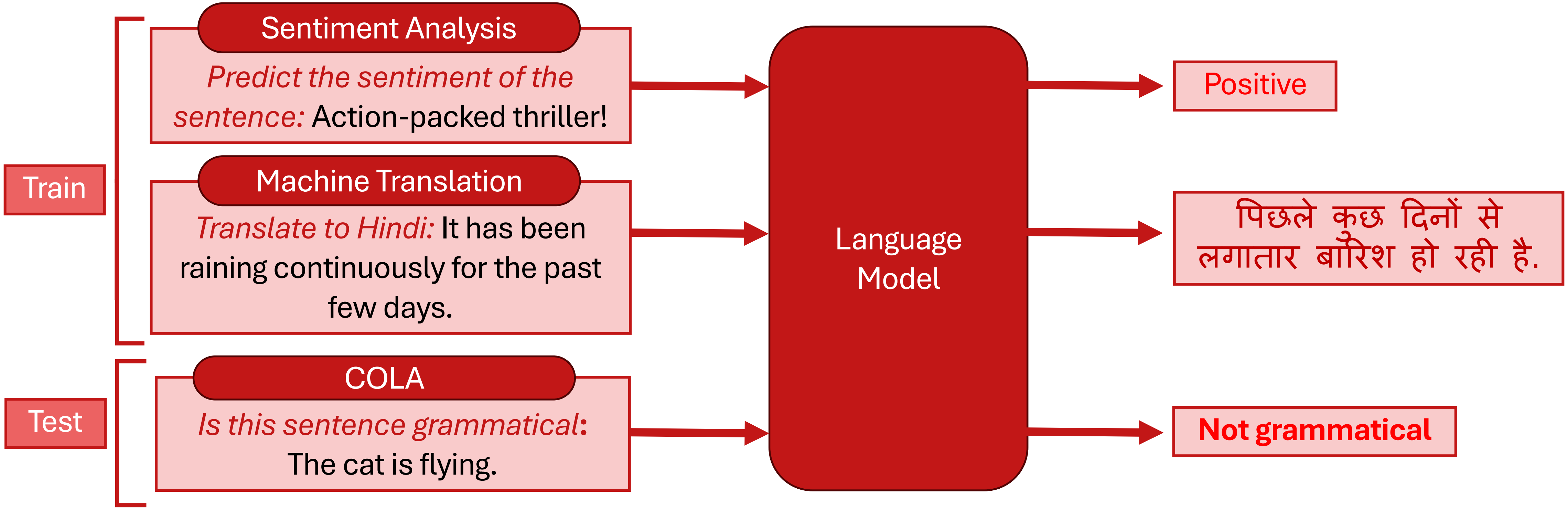
Classical multi-task learning aims to improve performance by combining different training tasks, allowing them to benefit from each other.



Problem: Model only generalizes within the training tasks and struggles with completely new tasks it hasn't encountered during training.

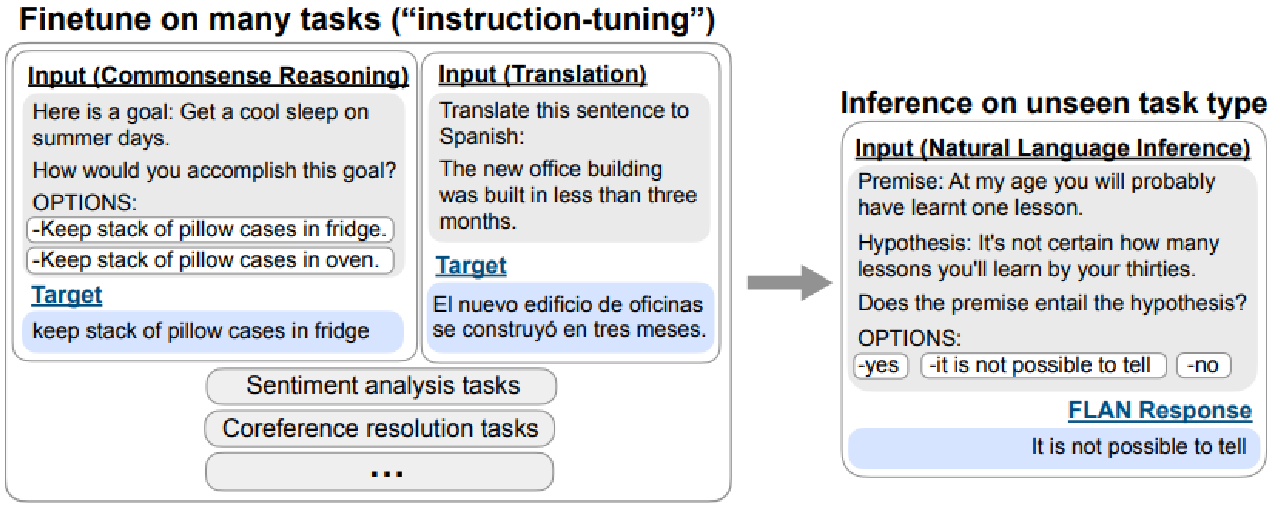
Solution – Describe task via instructions

Language models can generalize to new tasks by interpreting natural language instructions.



What is instruction tuning?

- Instruction tuning involves fine-tuning large language models on datasets described through natural language instructions.
  - Boosts their ability to understand and follow instructions and makes them better at handling new unseen tasks



Key Idea: By teaching (training) a language model to execute tasks based on instructions, it will learn to follow and apply them effectively to new, previously unseen tasks.

Paper: Wei, J., Bosma, M., Zhao, V.Y., Guu, K., Yu, A.W., Lester, B., Du, N., Dai, A.M. and Le, Q.V., 2021. Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*.

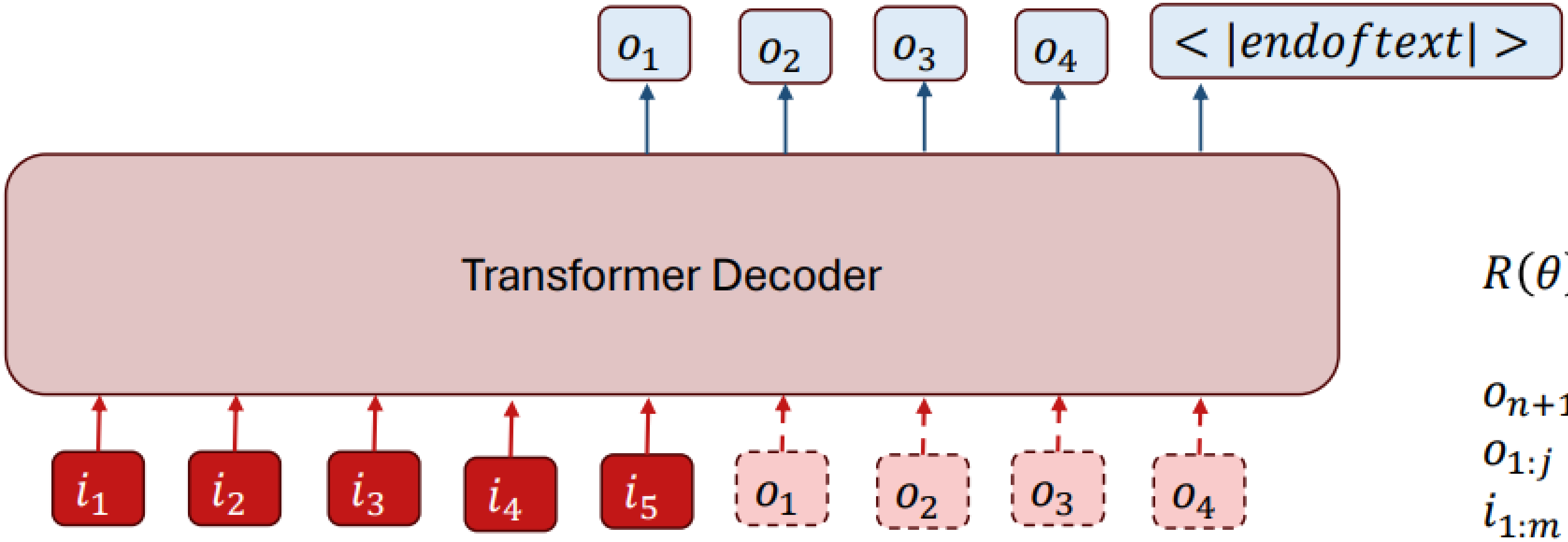
Training Loss

Prompt	Can you recommend some places to visit in Mumbai?
Response	Here are some great places you can explore in Mumbai ...

Input	Output
Can you recommend some places to visit in Mumbai?	Here
Can you recommend some places to visit in Mumbai? Here	are
Can you recommend some places to visit in Mumbai? Here are	some
Can you recommend some places to visit in Mumbai? Here are some	great
Can you recommend some places to visit in Mumbai? Here are some great	places
Can you recommend some places to visit in Mumbai? Here are some great places	you
...	...

Adapted from slides by Andrew Ng

Training Loss (Decoder-only models)



$$R(\theta) = \sum_{j=0}^{n} \log p_\theta(o_{j+1} | o_{1:j}, i_{1:m})$$

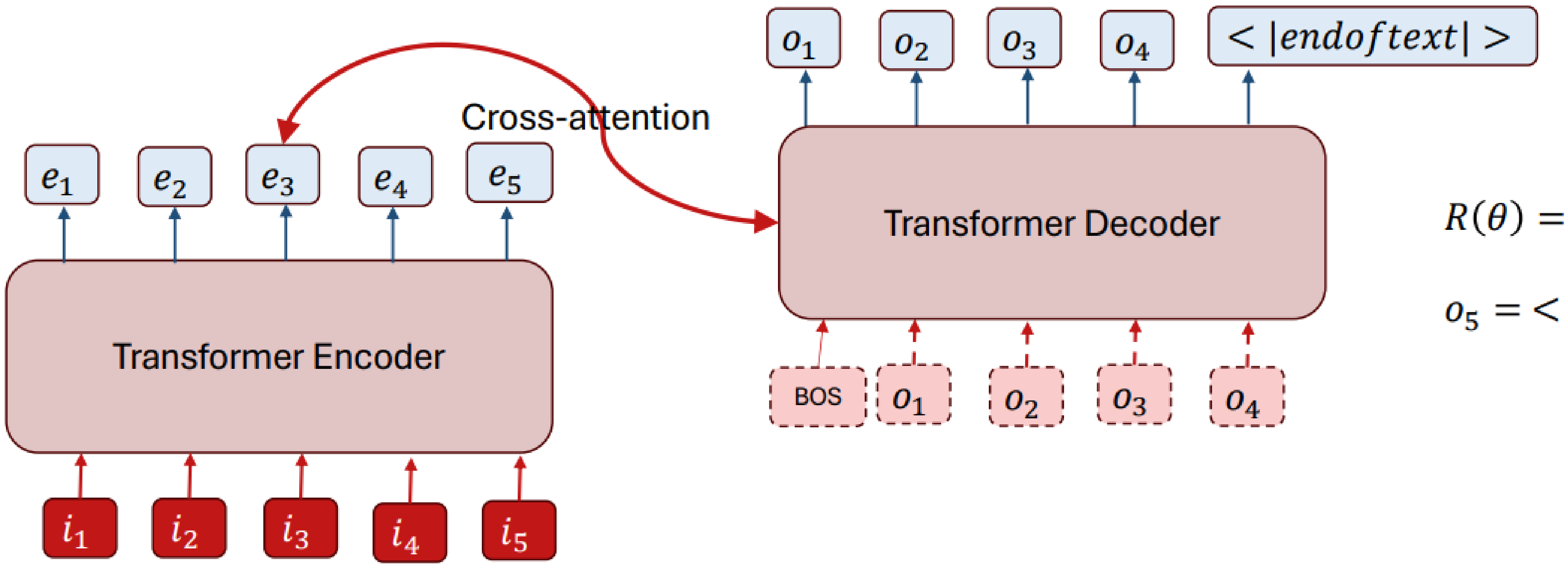
$$o_{n+1} = <|endoftext|>$$

$$o_{1:j} = o_1, \dots, o_j$$

$$i_{1:m} = i_1, \dots, i_m$$

Slide by Gaurav Pandey

Training Loss (Encoder-Decoder Models)



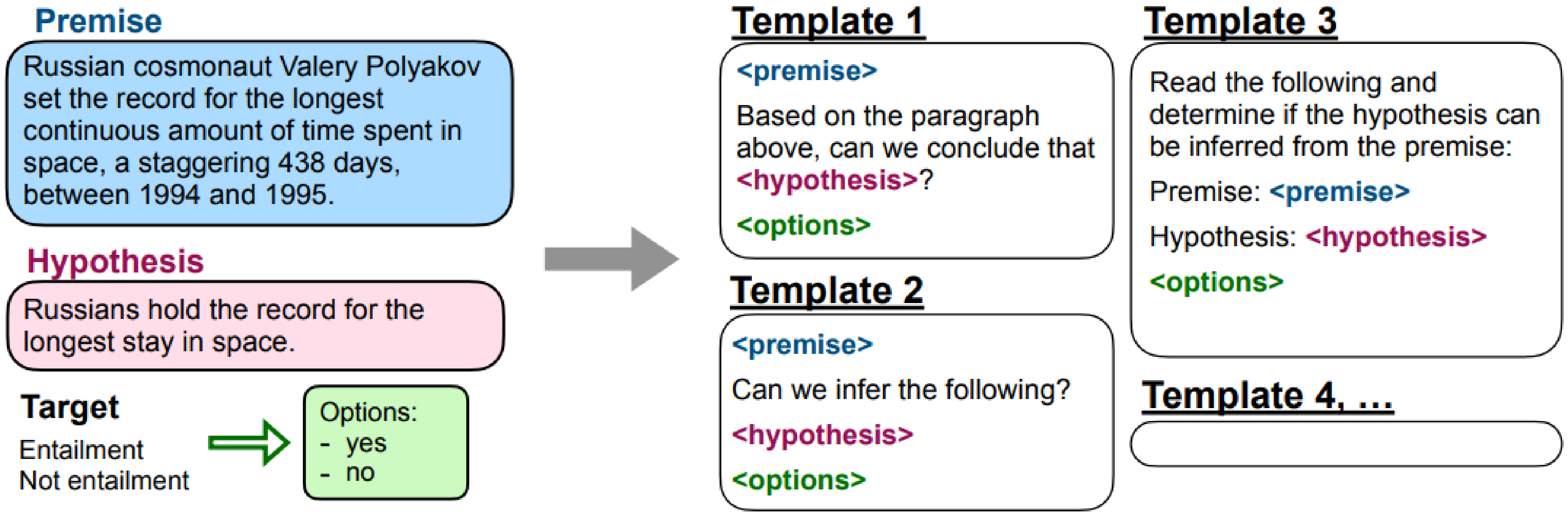
$$R(\theta) = \sum_{j=0}^{n} \log p_\theta(o_{j+1} | o_{1:j}, i_{1:m})$$

$$o_5 = <|endoftext|>$$

Slide by Gaurav Pandey

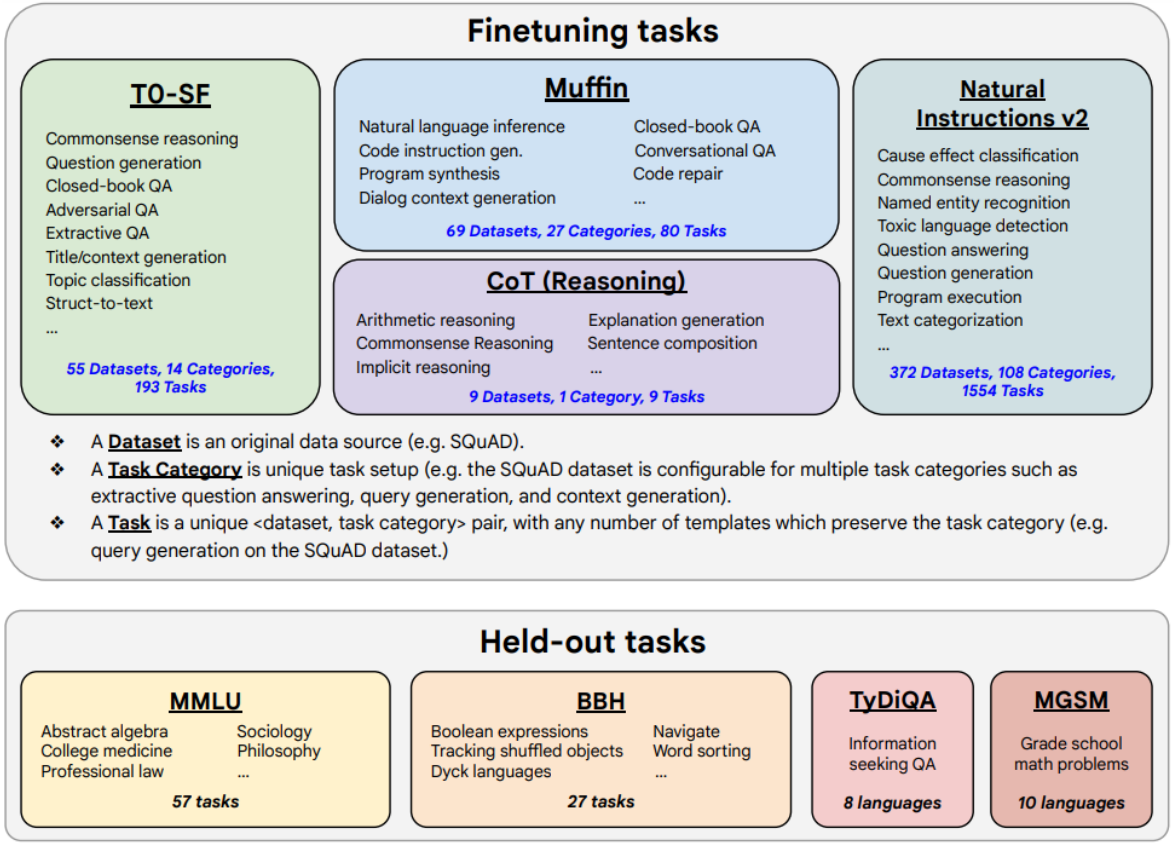
Multiple instruction templates for a task

Rephrasing the instructions for a task helps the model learn and generalize more effectively.



Paper: Wei, J., Bosma, M., Zhao, V.Y., Guu, K., Yu, A.W., Lester, B., Du, N., Dai, A.M. and Le, Q.V., 2021. Finetuned language models are zero-shot learners. *arXiv preprint arXiv:2109.01652*.

Diverse tasks



- The fine-tuning dataset consists of 473 datasets, 146 task categories, and 1,836 total tasks

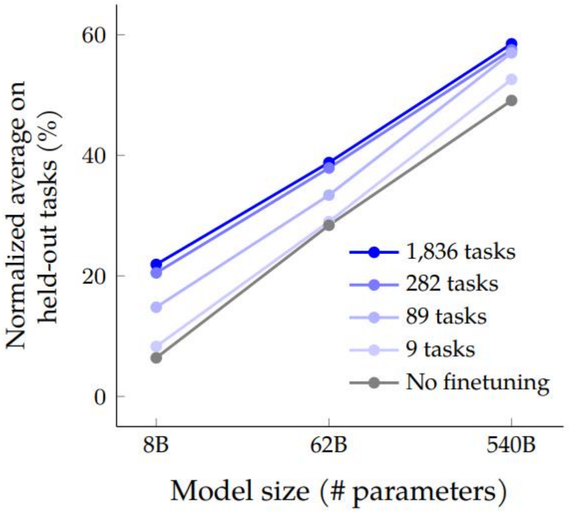
Paper: Chung, H.W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S. and Webson, A., 2024. Scaling instruction-finetuned language models. *Journal of Machine Learning Research*, *25*(70), pp.1-53.

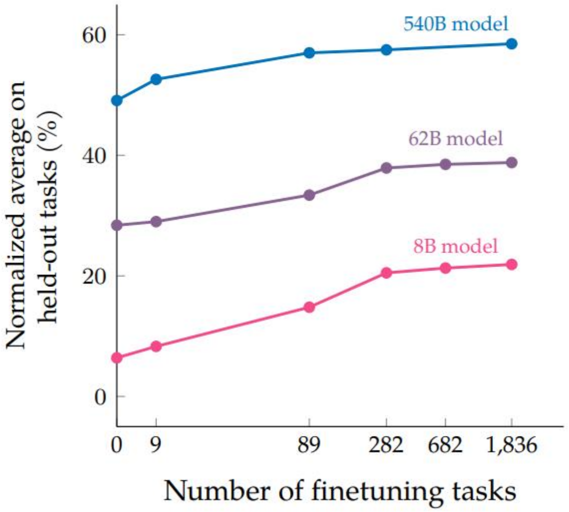
Fine-tuning data formats

	Without chain-of-thought	With chain-of-thought
Instruction without exemplars	Answer the following yes/no question. Can you write a whole Haiku in a single tweet? → yes	Answer the following yes/no question by reasoning step-by-step. Can you write a whole Haiku in a single tweet? → A haiku is a japanese three-line poem. That is short enough to fit in 280 characters. The answer is yes.
Instruction with exemplars	Q: Answer the following yes/no question. Could a dandelion suffer from hepatitis? A: no Q: Answer the following yes/no question. Can you write a whole Haiku in a single tweet? A: → yes	Q: Answer the following yes/no question by reasoning step-by-step. Could a dandelion suffer from hepatitis? A: Hepatitis only affects organisms with livers. Dandelions don't have a liver. The answer is no. Q: Answer the following yes/no question by reasoning step-by-step. Can you write a whole Haiku in a single tweet? A: → A haiku is a japanese three-line poem. That is short enough to fit in 280 characters. The answer is yes.

Paper: Chung, H.W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S. and Webson, A., 2024. Scaling instruction-finetuned language models. *Journal of Machine Learning Research*, *25*(70), pp.1-53.

Results



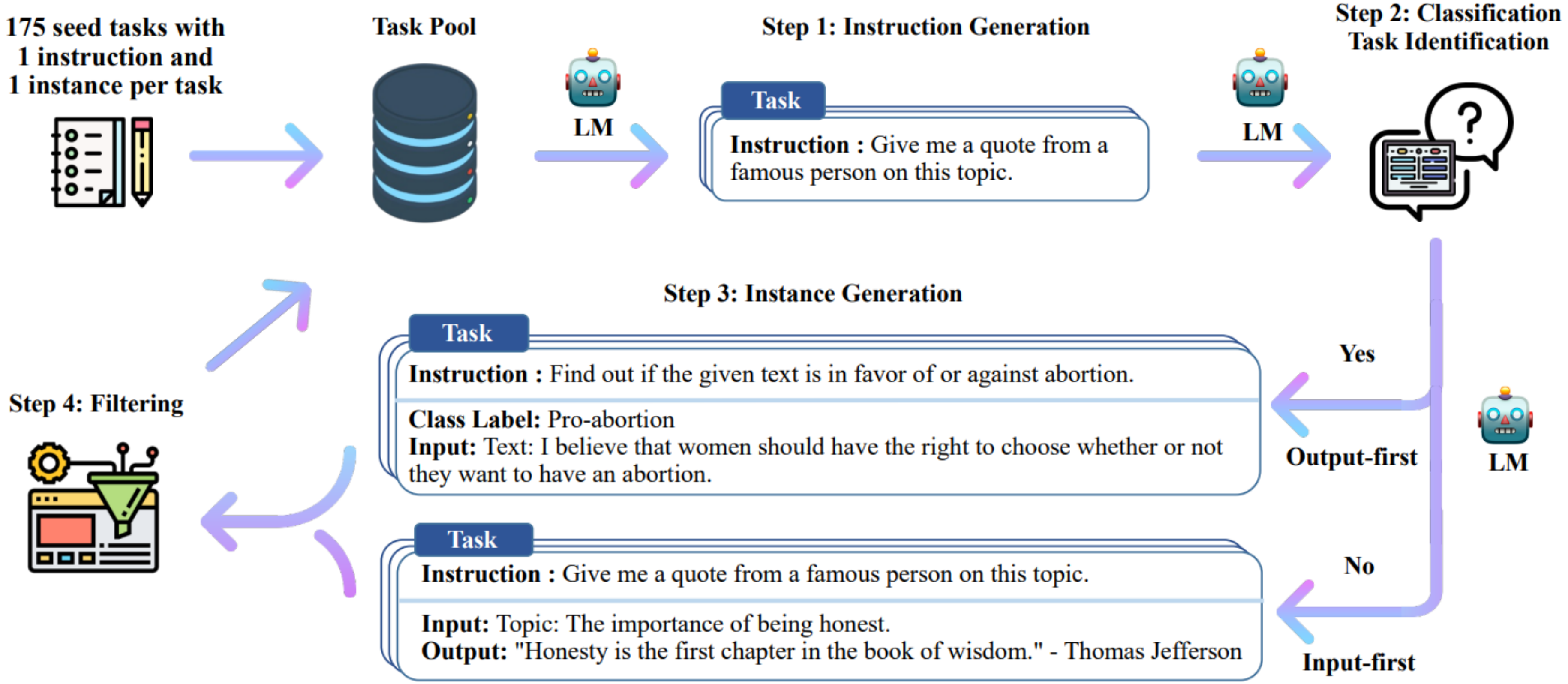


Params	Model	Norm. avg.	MMLU Direct	MMLU CoT	BBH Direct	BBH CoT	TyDiQA Direct	MGSM CoT
80M	T5-Small	-9.2	26.7	5.6	27.0	7.2	0.0	0.4
	Flan-T5-Small	-3.1 (+6.1)	28.7	12.1	29.1	19.2	1.1	0.2
250M	T5-Base	-5.1	25.7	14.5	27.8	14.6	0.0	0.5
	Flan-T5-Base	6.5 (+11.6)	35.9	33.7	31.3	27.9	4.1	0.4
780M	T5-Large	-5.0	25.1	15.0	27.7	16.1	0.0	0.3
	Flan-T5-Large	13.8 (+18.8)	45.1	40.5	37.5	31.5	12.3	0.7
3B	T5-XL	-4.1	25.7	14.5	27.4	19.2	0.0	0.8
	Flan-T5-XL	19.1 (+23.2)	52.4	45.5	41.0	35.2	16.6	1.9
11B	T5-XXL	-2.9	25.9	18.7	29.5	19.3	0.0	1.0
	Flan-T5-XXL	23.7 (+26.6)	55.1	48.6	45.3	41.4	19.0	4.9
8B	PaLM	6.4	24.3	24.1	30.8	30.1	25.0	3.4
	Flan-PaLM	21.9 (+15.5)	49.3	41.3	36.4	31.1	47.5	8.2
62B	PaLM	28.4	55.1	49.0	37.4	43.0	40.5	18.2
	Flan-PaLM	38.8 (+10.4)	59.6	56.9	47.5	44.9	58.7	28.5
540B	PaLM	49.1	71.3	62.9	49.1	63.7	52.9	45.9
	Flan-PaLM	58.4 (+9.3)	73.5	70.9	57.9	66.3	67.8	57.0
62B	cont-PaLM	38.1	61.2	57.6	41.7	53.1	45.7	32.0
	Flan-cont-PaLM	46.7 (+8.6)	66.1	62.0	51.0	53.3	62.7	40.3
540B	U-PaLM	50.2	71.5	64.0	49.2	62.4	54.6	49.9
	Flan-U-PaLM	59.1 (+8.9)	74.1	69.8	59.3	64.9	68.3	60.4

Params	Model	Architecture	Pre-training Objective	Pre-train FLOPs	Finetune FLOPs	% Finetune Compute
80M	Flan-T5-Small	encoder-decoder	span corruption	1.8E+20	2.9E+18	1.6%
250M	Flan-T5-Base	encoder-decoder	span corruption	6.6E+20	9.1E+18	1.4%
780M	Flan-T5-Large	encoder-decoder	span corruption	2.3E+21	2.4E+19	1.1%
3B	Flan-T5-XL	encoder-decoder	span corruption	9.0E+21	5.6E+19	0.6%
11B	Flan-T5-XXL	encoder-decoder	span corruption	3.3E+22	7.6E+19	0.2%
8B	Flan-PaLM	decoder-only	causal LM	3.7E+22	1.6E+20	0.4%
62B	Flan-PaLM	decoder-only	causal LM	2.9E+23	1.2E+21	0.4%
540B	Flan-PaLM	decoder-only	causal LM	2.5E+24	5.6E+21	0.2%
62B	Flan-cont-PaLM	decoder-only	causal LM	4.8E+23	1.8E+21	0.4%
540B	Flan-U-PaLM	decoder-only	prefix LM + span corruption	2.5E+23	5.6E+21	0.2%

Paper: Chung, H.W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S. and Webson, A., 2024. Scaling instruction-finetuned language models. *Journal of Machine Learning Research*, *25*(70), pp.1-53.

Self-Instruct



Paper: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. [Self-Instruct: Aligning Language Models with Self-Generated Instructions](). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada. Association for Computational Linguistics.

Prompt Templates

```
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:
```

New instruction generation

```
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:  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}
```

Classify a task instruction is a classification task or not

Paper: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. [Self-Instruct: Aligning Language Models with Self-Generated Instructions]. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada. Association for Computational Linguistics.

Prompt Templates

```
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: 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}
```

Input-first approach of instance generation

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

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

Task: 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: 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}
```

Output-first approach of instance generation

Paper: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. [Self-Instruct: Aligning Language Models with Self-Generated Instructions](). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada. Association for Computational Linguistics.

Filtering

- A new instruction is included in the task pool only if its ROUGE-L similarity score with any existing instruction is less than 0.7.
- Exclude instructions that have certain keywords such as image, picture, graph
- Remove duplicate instances or ones with the same input but different outputs.
- Filtering based on heuristics (e.g., length of the instruction).

Examples of generated data

Instruction + Instance Input	Target Output

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: Write a paragraph describing how to tie shoelaces for people who have never tied their shoes before. explain each step briefly.	Paragraph: To tie shoelaces, first you need to make a loop with one end of the lace. Then pass the other end through the loop and pull it tight. Next, you need to wrap the lace around the shoe several times. Finally, tuck in the loose end into the wrapped part of the lace.
Instruction: Given a word, find out its length and its number of vowels. Input: Word = "hello"	Length = 5, Number of vowels = 2
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.

Paper: Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. Self-Instruct: Aligning Language Models with Self-Generated Instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 13484–13508, Toronto, Canada. Association for Computational Linguistics.

Various Instruction Tuning Datasets

Type	Dataset Name	# of Instances	# of Lang	Construction	Open-source
Human-Crafted	UnifiedQA (Khashabi et al., 2020)[1]	750K	En	human-crafted	Yes
	UnifiedSKG (Xie et al., 2022)[3]	0.8M	En	human-crafted	Yes
	Natural Instructions (Honovich et al., 2022)[4]	193K	En	human-crafted	Yes
	Super-Natural Instructions (Wang et al., 2022f)[5]	5M	55 Lang	human-crafted	Yes
	P3 (Sanh et al., 2021)[6]	12M	En	human-crafted	Yes
	xP3 (Muennighoff et al., 2022)[7]	81M	46 Lang	human-crafted	Yes
	Flan 2021 (Longpre et al., 2023)[8]	4.4M	En	human-crafted	Yes
	COIG (Zhang et al., 2023a)[9]	-	-	-	Yes
	InstructGPT (Ouyang et al., 2022)	13K	Multi	human-crafted	No
	Dolly (Conover et al., 2023a)[16]	15K	En	human-crafted	Yes
	LIMA (Zhou et al., 2023a)[18]	1K	En	human-crafted	Yes
	ChatGPT (OpenAI, 2022)	-	Multi	human-crafted	No
	OpenAssistant (Köpf et al., 2023)[20]	161,443	Multi	human-crafted	Yes

Synthetic Data (Distillation)	OIG (LAION.ai, 2023)[2]	43M	En	ChatGPT (No technique reports)	Yes
	Unnatural Instructions (Honovich et al., 2022)[10]	240K	En	InstructGPT-Generated	Yes
	InstructWild (Xue et al., 2023)[12]	104K	-	ChatGPT-Generated	Yes
	Evol-Instruct / WizardLM (Xu et al., 2023a)[13]	52K	En	ChatGPT-generated	Yes
	Alpaca (Taori et al., 2023a)[14]	52K	En	InstructGPT-generated	Yes
	LogiCoT (Liu et al., 2023a)[15]	-	En	GPT-4-Generated	Yes
	GPT-4-LLM (Peng et al., 2023)[17]	52K	En&Zh	GPT-4-Generated	Yes
	Vicuna (Chiang et al., 2023)	70K	En	Real User-ChatGPT Conversations	No
	Baize v1 (Conover et al., 2023b)[21]	111.5K	En	ChatGPT-Generated	Yes
	UltraChat (Ding et al., 2023a)[22]	675K	En&Zh	GPT 3/4-Generated	Yes
	Guanaco (JosephusCheung, 2021)[19]	534,530	Multi	GPT (Unknown Version)-Generated	Yes
	Orca (Mukherjee et al., 2023)[23]	1.5M	En	GPT 3.5/4-Generated	Yes
	ShareGPT[24]	90K	Multi	Real User-ChatGPT Conversations	Yes
	WildChat[25]	150K	Multi	Real User-ChatGPT Conversations	Yes
	WizardCoder (Luo et al., 2023)	-	Code	LLaMa 2-Generated	No
	Magicoder (Wei et al., 2023b)[26]	75K/110K	Code	GPT-3.5-Generated	Yes
	WaveCoder (Yu et al., 2023)	-	Code	GPT 4-Generated	No
	Phi-1 (Gunasekar et al., 2023)[27]	6B Tokens	Code Q and A	GPT-3.5-Generated	Yes
	Phi-1.5 (Li et al., 2023i)	-	Code Q and A	GPT-3.5-Generated	No
	Nectar (Zhu et al., 2023a)[28]	183K	En	GPT 4-Generated	Yes
Synthetic Data (Self-Improvement)	Self-Instruct (Wang et al., 2022c)[11]	52K	En	InstructGPT-Generated	Yes
	Instruction Backtranslation (Li et al., 2023g)	502K	En	LLaMa-Generated	No
	SPIN (Chen et al., 2024b)[29]	49.8K	En	Zephyr-Generated	Yes

Paper: Zhang, S., Dong, L., Li, X., Zhang, S., Sun, X., Wang, S., Li, J., Hu, R., Zhang, T., Wu, F. and Wang, G., 2023. Instruction tuning for large language models: A survey. *arXiv preprint arXiv:2308.10792*.