大模型论文汇报

王永胜

论文

- Large Language Models Are Reasoning Teachers--2023ACL
- Check Your Facts and Try Again: Improving Large Language Models with External Knowledge and Automated Feedback--2023arxiv.org

Large Language Models Are Reasoning Teachers--2023ACL

1、解决的问题：大模型计算成本高，大模型作为推理教师使得小模型可实现复杂推理。

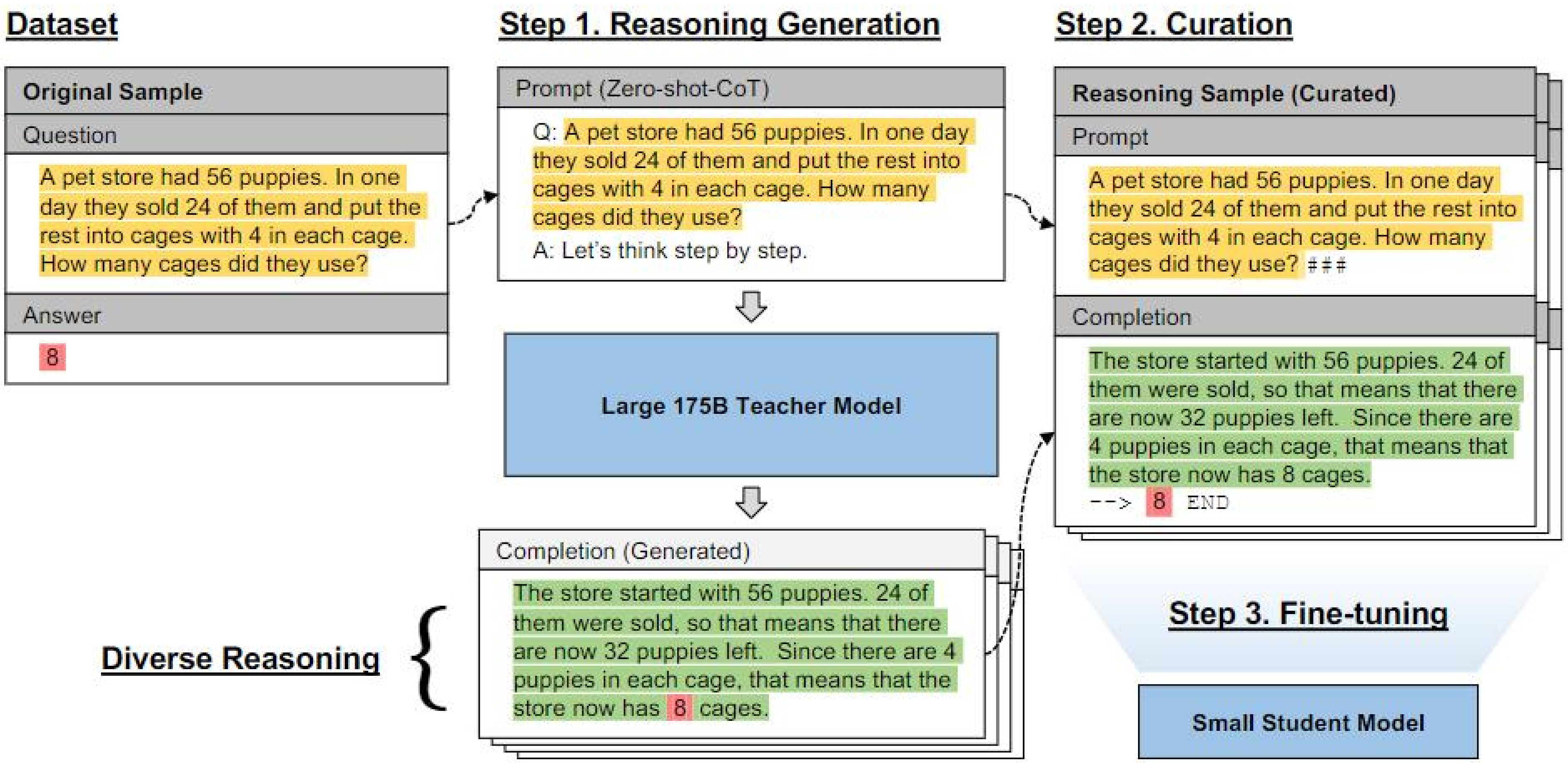
2、创新点：提出Fine-tune-CoT（利用大模型的推理能力教会小模型如何解决复杂任务），在没有任何人工标注的情况下提高了学生模型的性能。

3、方法：

1）Reasoning generation（"Q: <qi>. A: Let's think step by step. <^ri> Therefore, the answer is <^ai>".）

2）Curation

3）Fine-tune


Dataset
Original Sample
Question
A pet store had 56 puppies. In one day they sold 24 of them and put the rest into cages with 4 in each cage. How many cages did they use?
Answer
8
Step 1. Reasoning Generation
Prompt (Zero-shot-CoT)
Q: A pet store had 56 puppies. In one day they sold 24 of them and put the rest into cages with 4 in each cage. How many cages did they use?
A: Let's think step by step.
Large 175B Teacher Model
Completion (Generated)
The store started with 56 puppies. 24 of them were sold, so that means that there are now 32 puppies left. Since there are 4 puppies in each cage, that means that the store now has 8 cages.
Diverse Reasoning
Step 2. Curation
Reasoning Sample (Curated)
Prompt
A pet store had 56 puppies. In one day they sold 24 of them and put the rest into cages with 4 in each cage. How many cages did they use? ###
Completion
The store started with 56 puppies. 24 of them were sold, so that means that there are now 32 puppies left. Since there are 4 puppies in each cage, that means that the store now has 8 cages.
--> 8 END
Step 3. Fine-tuning
Small Student Model

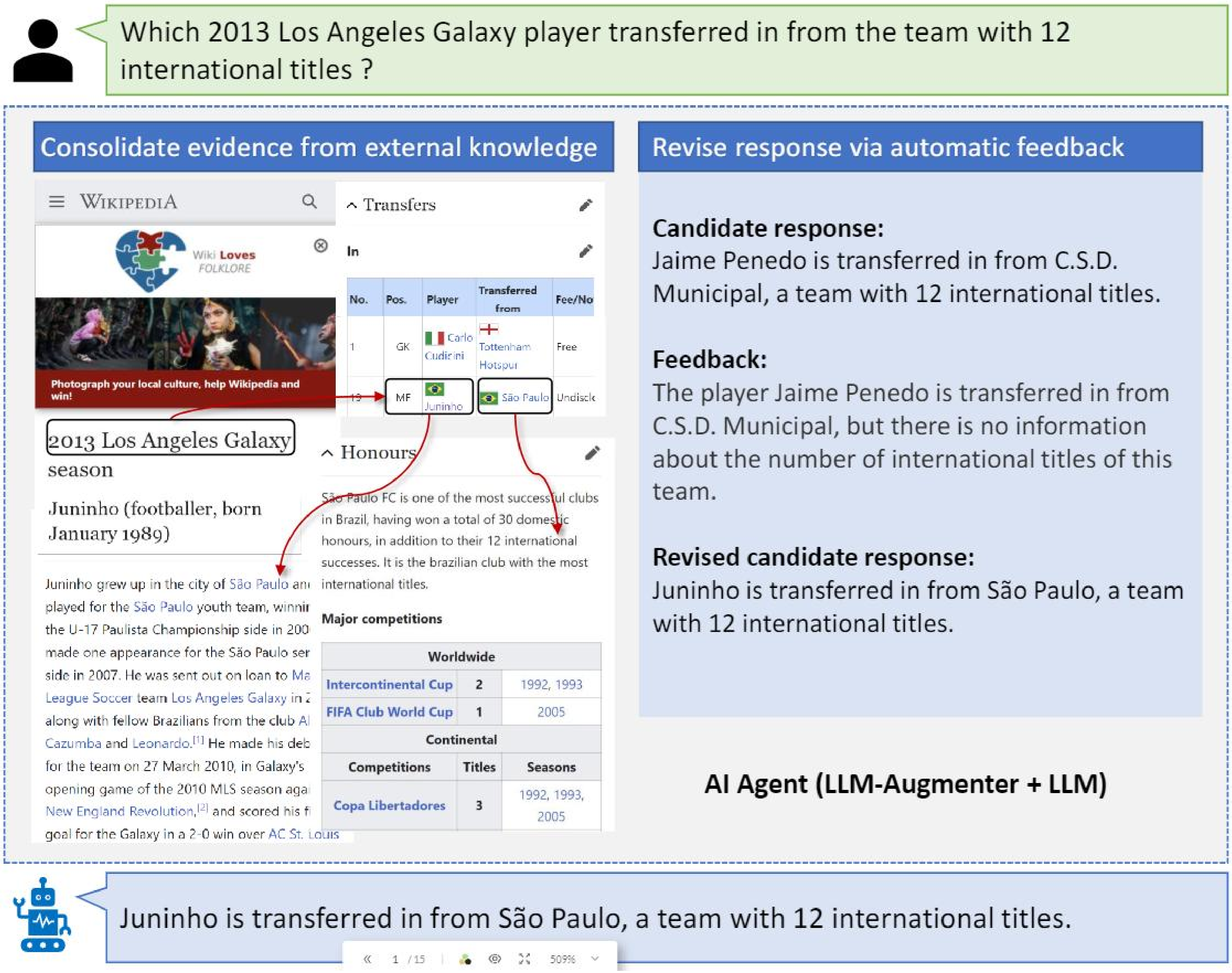
Method	Params	Single Eq	Add Sub	Multi Arith	GSM8K	Aqua	SVAMP	Date Understanding	Shuffled Objects	Last Letter	Coin Flip	Common SenseQA	Strategy QA
Random		0.00	0.00	0.00	0.00	20.00	0.00	17.12	33.33	0.00	50.00	20.00	50.00
Teacher: InstructGPT (`text-davinci-002`)													
Zero-shot-CoT	175B	82.24	78.99	78.89	40.26	34.25	64.67	73.87	50.22	56.00	92.67	61.75	53.57
Student: GPT-3 (`ada`, `babbage`, `curie`)													
Zero-shot	6.7B	0.66	0.84	3.33	1.74	16.54	2.67	9.91	32.89	0.00	56.67	20.23	52.98
Zero-shot-CoT	6.7B	1.32	2.52	5.00	2.35	21.26	1.33	15.32	31.11	0.00	46.67	19.98	51.09
Few-shot-CoT	6.7B	22.37	**31.93**	10.00	2.50	15.75	11.33	12.84	-	0.67	40.00	24.73	54.68
Fine-tune	6.7B	**24.34**	25.21	15.00	6.14	15.35	20.67	14.41	33.78	32.67	72.00	**76.17**	**65.21**
Fine-tune-CoT	0.3B	7.24	6.72	6.11	3.11	23.62	5.00	17.12	49.33	50.67	99.33	32.68	52.55
	1.3B	11.18	11.76	13.33	4.70	19.69	8.00	38.74	52.44	50.67	100.00	43.08	52.69
	6.7B	20.39	21.01	33.33	**6.75**	**24.02**	12.67	60.36	64.44	52.67	98.67	56.76	55.02
Fine-tune-CoT	0.3B	9.21	10.08	23.89	-	-	14.33	58.56	61.78	59.33	99.33	-	57.21
w/ diverse reasoning	1.3B	18.42	19.33	27.78	-	-	16.33	70.27	72.00	60.67	100.00	-	57.06
	6.7B	**24.34**	31.09	**53.33**	-	-	**30.33**	**83.78**	**73.33**	**62.00**	**100.00**	-	58.22

Table 1: **Fine-tune-CoT Performance.** Accuracy (%) of OpenAI models on 12 tasks under Fine-tune-CoT (with diverse reasoning) and baseline methods. 'Random' refers to random-guess performance derived based on the number of choices in multi-choice tasks. For diverse reasoning, we report results for maximum degree D considered: $D = 64$ for MultiArith and SVAMP; $D = 8$ for other datasets. We omit diverse reasoning for large datasets due to resource constraints and Few-shot-CoT for Tracking Shuffled Objects due to absence of prompts.

- 分析（消融实验、开放域）
- 讨论（可行性，易实现性，原因，展望）

Check Your Facts and Try Again: Improving Large Language Models with External Knowledge and Automated Feedback--2023arxiv.org

- 解决的问题：大模型对于有些知识类的问题比如实时新闻，私域知识等，容易产生幻觉的问题。
- 创新点：除了引入外部知识，还进行迭代反馈，提高准确性。
- 方法：以现有大模型为基座，引入外部知识和自动反馈机制。


Which 2013 Los Angeles Galaxy player transferred in from the team with 12 international titles ?
Consolidate evidence from external knowledge
Revise response via automatic feedback
Wikipedia
Wiki Loves Folklore
Photograph your local culture, help Wikipedia and win!
2013 Los Angeles Galaxy season
Juninho (footballer, born January 1989)
Juninho grew up in the city of São Paulo an played for the São Paulo youth team, winnir the U-17 Paulista Championship side in 200 made one appearance for the São Paulo ser side in 2007. He was sent out on loan to Ma League Soccer team Los Angeles Galaxy in 2 along with fellow Brazilians from the club Al Cazumba and Leonardo.[1] He made his deb for the team on 27 March 2010, in Galaxy's opening game of the 2010 MLS season agai New England Revolution,[2] and scored his fi goal for the Galaxy in a 2-0 win over AC St. Louis
Transfers
In
No. | Pos. | Player | Transferred from | Fee/No
1 | GK | Carlo Cudicini | Tottenham Hotspur | Free
13 | MF | Juninho | São Paulo | Undiscl
Honours
São Paulo FC is one of the most successful clubs in Brazil, having won a total of 30 domestic honours, in addition to their 12 international successes. It is the brazilian club with the most international titles.
Major competitions
Worldwide
Intercontinental Cup | 2 | 1992, 1993
FIFA Club World Cup | 1 | 2005
Continental
Competitions | Titles | Seasons
Copa Libertadores | 3 | 1992, 1993, 2005
Candidate response:
Jaime Penedo is transferred in from C.S.D. Municipal, a team with 12 international titles.
Feedback:
The player Jaime Penedo is transferred in from C.S.D. Municipal, but there is no information about the number of international titles of this team.
Revised candidate response:
Juninho is transferred in from São Paulo, a team with 12 international titles.
AI Agent (LLM-Augmenter + LLM)
Juninho is transferred in from São Paulo, a team with 12 international titles.
1 / 15
509%

LLM-AUGMENTER架构

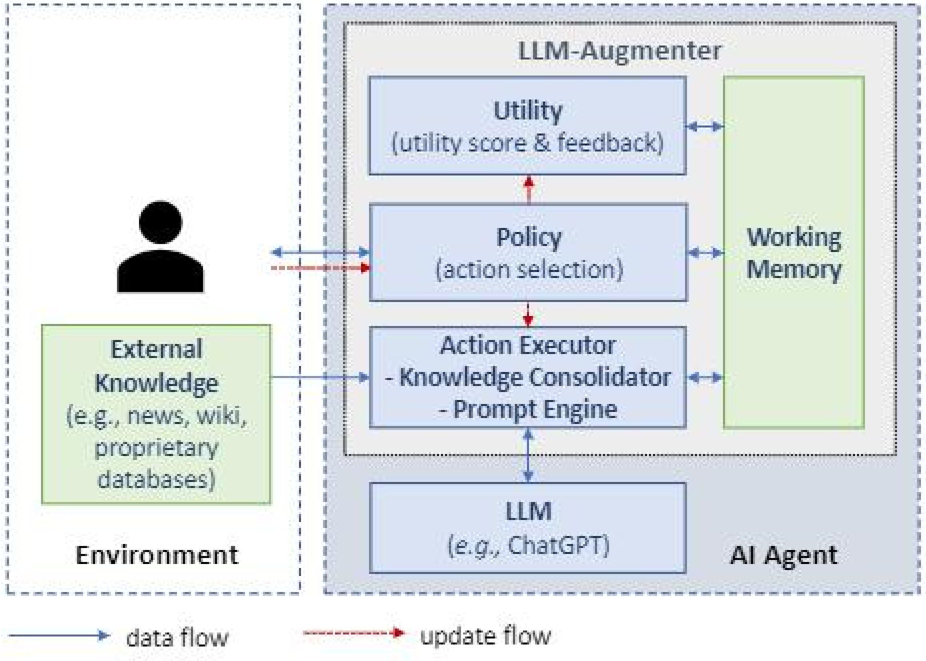


Figure 2: LLM-AUGMENTER architecture showing how its plug-and-play modules interact with the LLM and the user's environment.

Prompt Engine：根据任务，设计Prompt将知识、上下文，候选等信息融合起来。

Utility: 使用ChatGPT来根据任务关注的维度（知识、回复候选）来做正确性分值和问题反馈。

Knowledge Consolidator: 一个基于BM25的检索器，根据任务去检索Web知识，或者本地FAQ知识库。

Policy:基于任务设计Rule，训练一个策略模型Pi，使得当前状态下选择出的action，期望reward最大。T5

知识整合器和自动反馈机制在两个任务上的效果

Model	K.C.	Feedback	KF1 ↑	BLEU ↑	ROUGE ↑	chrF ↑	METEOR ↑	BERTScore ↑	BARTScore ↑	BLEURT ↑	Avg. length
CHATGPT	-	-	26.71	1.01	16.78	23.80	7.34	82.14	0.25	26.98	58.94
LLM-AUGMENTER	BM25	✗	34.96	6.71	22.25	27.02	9.35	**83.46**	0.34	26.89	46.74
LLM-AUGMENTER	BM25	✓	**36.41**	**7.63**	**22.80**	**28.66**	**10.17**	83.33	**0.35**	**27.71**	54.24
LLM-AUGMENTER	gold	✗	*57.44*	*19.24*	*38.89*	*40.02*	*17.21*	*86.65*	*0.82*	*40.55*	*44.35*
LLM-AUGMENTER	gold	✓	*60.76*	*21.49*	*40.56*	*42.14*	*18.50*	*86.89*	*0.93*	*42.15*	*47.19*

Table 1: Evaluation scores (in %) and average response lengths for the News Chat (DSTC7) dataset. BM25: Each model retrieves 5 knowledge snippets from the corresponding knowledge source. K.C. denotes Knowledge Consolidator.

Model	K.C.	Feedback	KF1 ↑	BLEU ↑	ROUGE ↑	chrF ↑	METEOR ↑	BERTScore ↑	BARTScore ↑	BLEURT ↑	Avg. length
CHATGPT	-	-	31.33	4.70	24.02	27.14	12.83	87.88	1.53	**47.99**	28.81
LLM-AUGMENTER	BM25	✗	34.07	**4.78**	**24.52**	28.95	13.61	**87.96**	1.78	47.21	32.65
LLM-AUGMENTER	BM25	✓	**37.41**	3.86	24.20	**30.90**	**14.74**	87.58	**2.09**	44.71	45.07
LLM-AUGMENTER	gold	✗	*45.63*	*6.54*	*29.77*	*33.32*	*16.93*	*89.35*	*2.59*	*54.38*	*33.04*
LLM-AUGMENTER	gold	✓	*52.83*	*5.63*	*29.65*	*35.68*	*18.66*	*89.01*	*3.14*	*52.49*	*45.09*

Table 2: Evaluation scores (in %) and average response lengths for the Customer Service (DSTC11) dataset. BM25: Each model retrieves 5 knowledge snippets from the corresponding knowledge source. K.C. denotes Knowledge Consolidator.

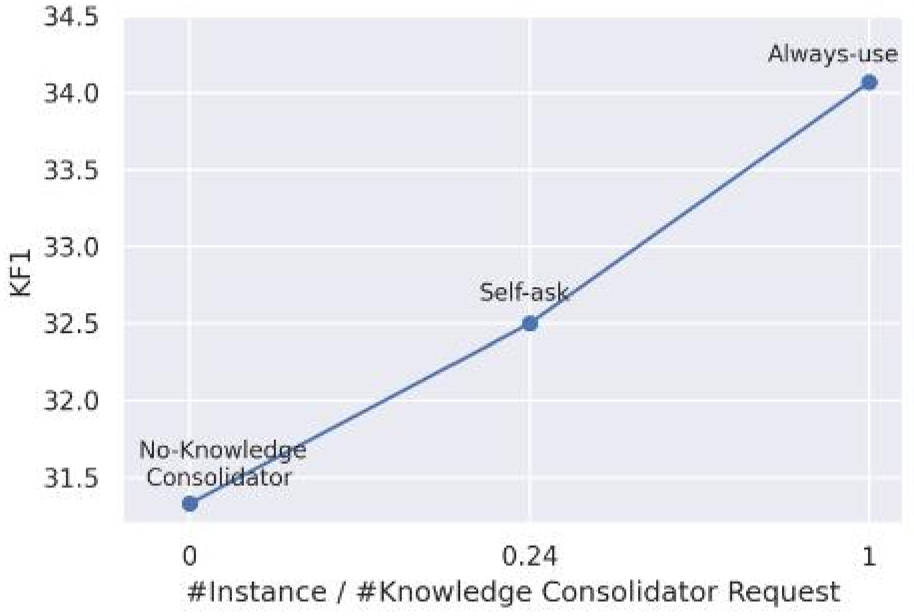


Figure 4: Ablation studies on different policies of LLM-AUGMENTER in Customer Service scenario.

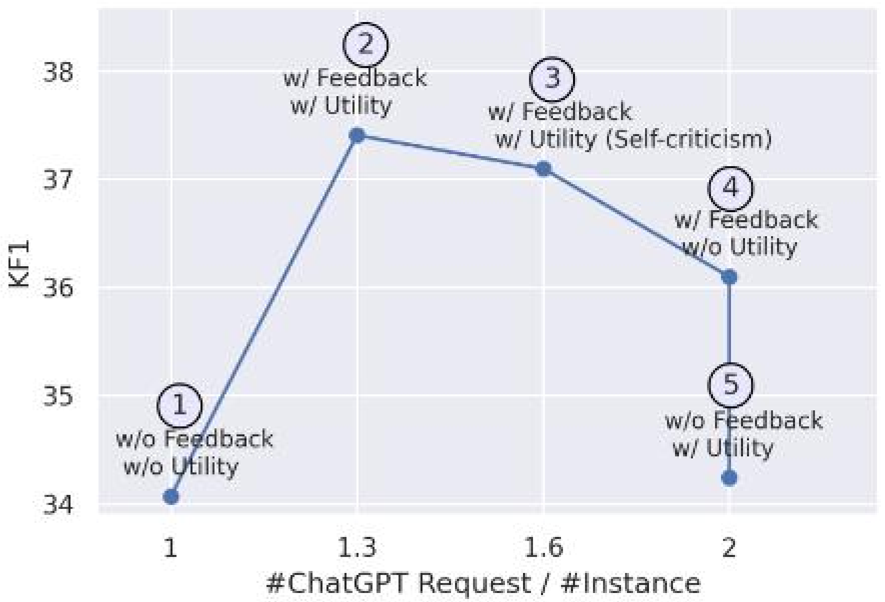


Figure 5: LLM-AUGMENTER benefits from the combination of using utility function and iterative improvement with feedback. The x-axis indicates the average number of ChatGPT prompting and the y-axis is the KF1. The studies are conducted in the Customer Service scenario with knowledge being provided by BM25.