

KCC 2023

Self-Instruct 데이터셋을 활용한 언어 모델 변화에 따른 한국어 지시학습 성능 비교

이현민*, 나승훈, 임준호, 김태형, 류휘정, 장두성
전북대학교*, 한국전자통신연구원, KT
leehm@jbnu.ac.kr*



목차

1. 연구 개요
2. 관련 연구
3. 한국어 지시학습(Instruct Tuning)
 - 사전학습 언어모델
 - 지시학습 데이터 셋
 - 실험 세팅
4. 실험 결과
 - 실험 결과 및 분석

1. 연구 개요: Instruct Tuning

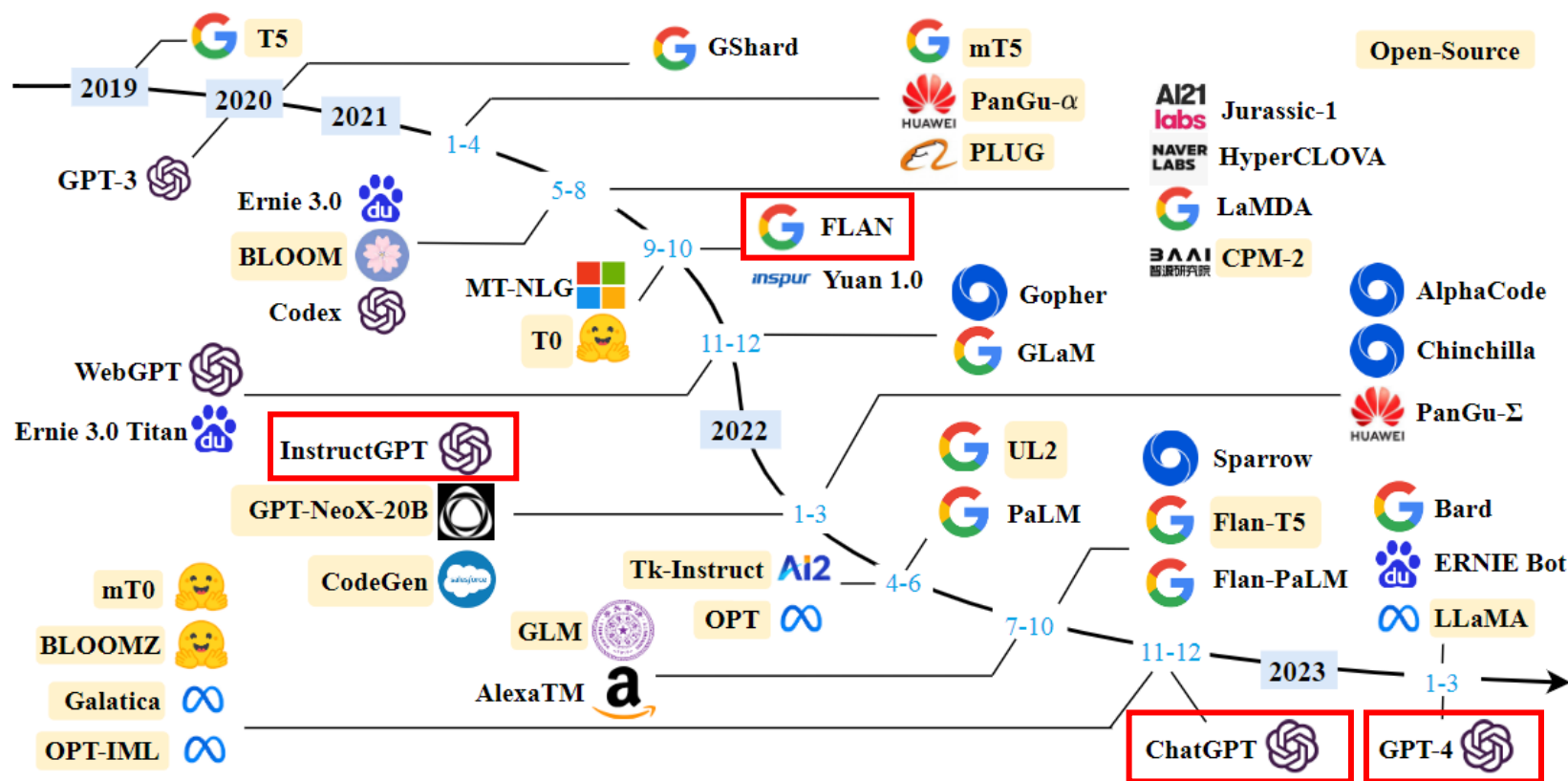
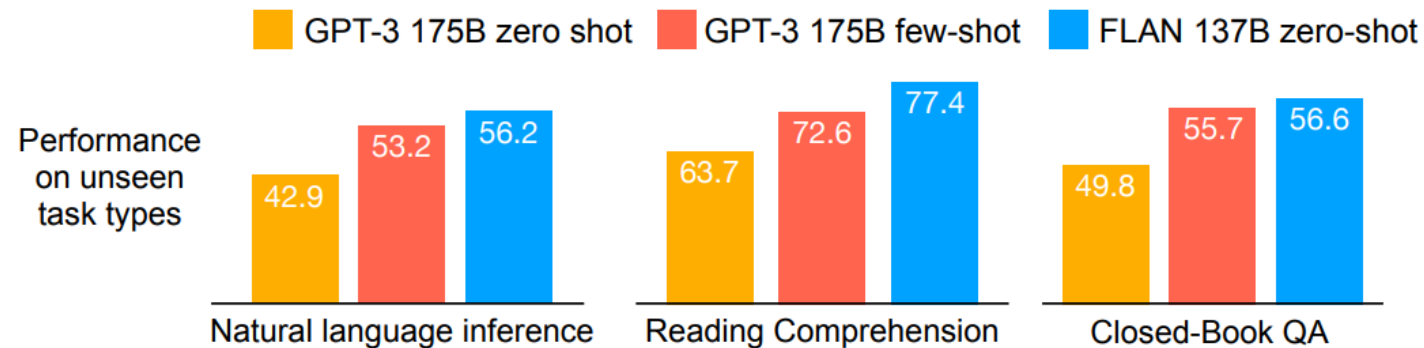
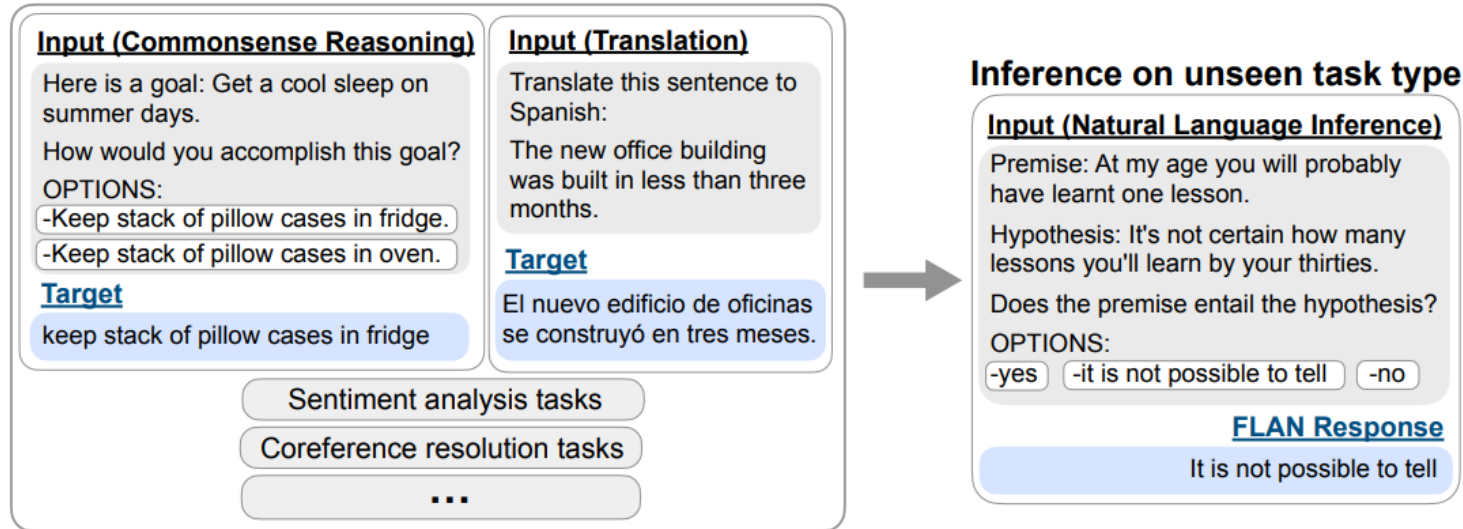


Fig. 1. A timeline of existing large language models (having a size larger than 10B) in recent years. We mark the open-source LLMs in yellow color.

2. 관련 연구: Finetuned language models are zero-shot learners (FLAN)

Finetune on many tasks (“instruction-tuning”)



2. 관련 연구: Training language models to follow instructions with human feedback

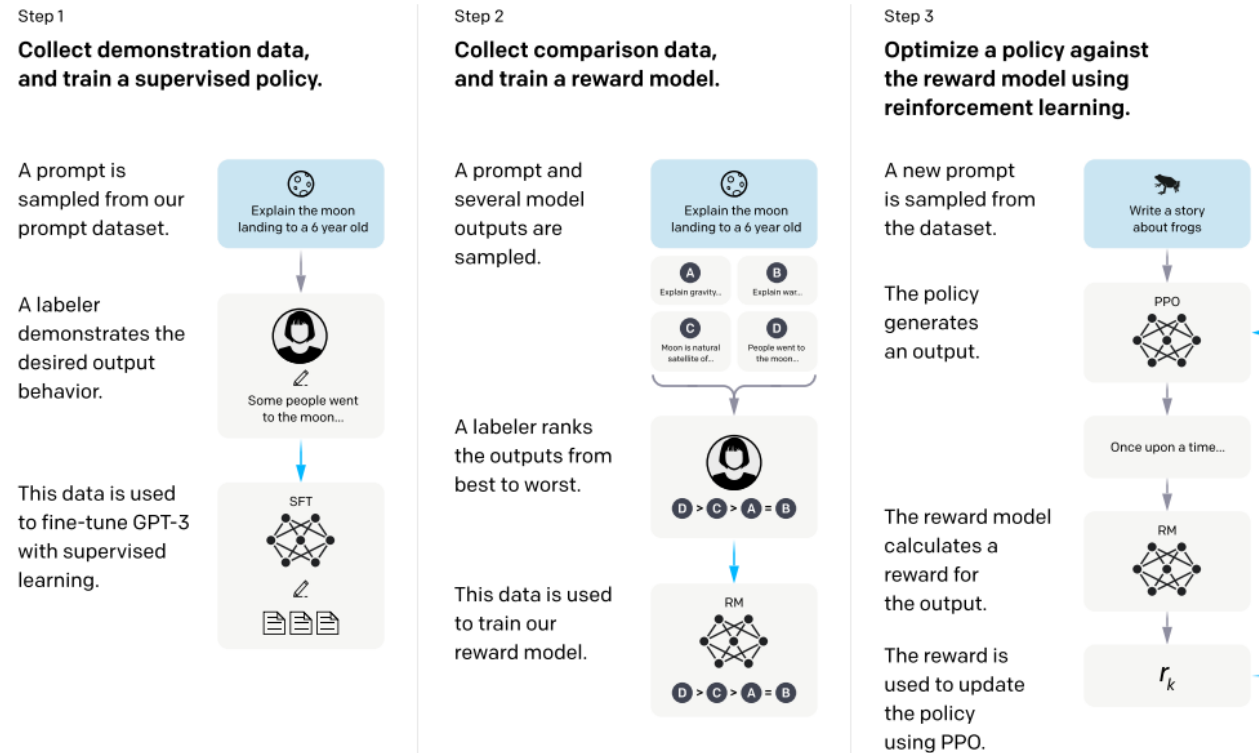


Figure 2: A diagram illustrating the three steps of our method: (1) supervised fine-tuning (SFT), (2) reward model (RM) training, and (3) reinforcement learning via proximal policy optimization (PPO) on this reward model. Blue arrows indicate that this data is used to train one of our models. In Step 2, boxes A-D are samples from our models that get ranked by labelers. See Section 3 for more details on our method.

2. 관련 연구: Self-Instruct: Aligning Language Models with Self-Generated Instructions

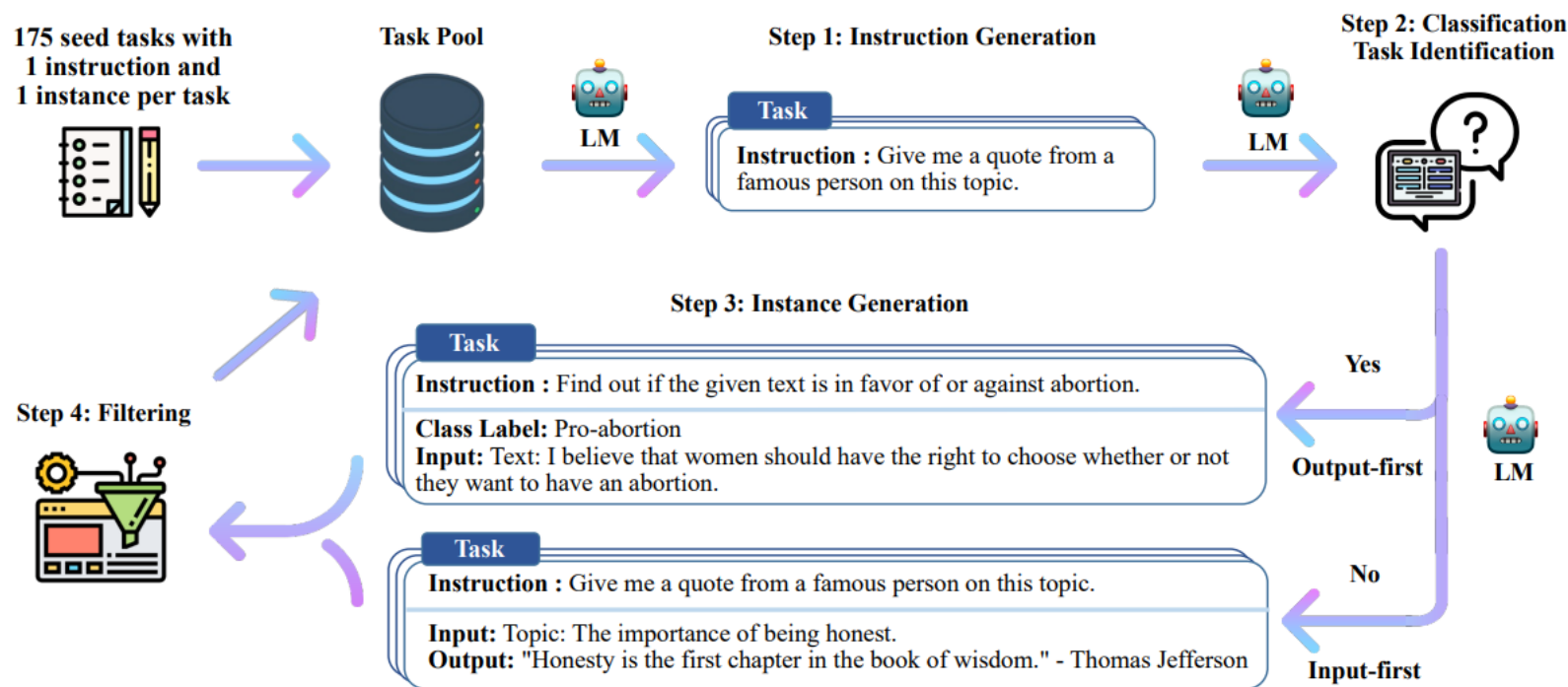


Figure 2: A high-level overview of SELF-INSTRUCT. The process starts with a small seed set of tasks as the task pool. Random tasks are sampled from the task pool, and used to prompt an off-the-shelf LM to generate both new instructions and corresponding instances, followed by filtering low-quality or 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.

3. 한국어 지시학습: 사전학습 언어모델

Hyperparameter	1.3B	3.8B	5.8B	12.8B
$n_{parameters}$	1,331,810,304	3,809,974,272	5,885,059,072	12,898,631,680
n_{layers}	24	32	28	40
d_{model}	2,048	3,072	4,096	5,120
d_{ff}	8,192	12,288	16,384	20,480
n_{heads}	16	24	16	40
d_{head}	128	128	256	128
n_{vocab}	30,003 / 30,080	30,003 / 30,080	30,003 / 30,080	30,003 / 30,080
Positional Encoding	Rotary (RoPE)	Rotary (RoPE)	Rotary (RoPE)	Rotary (RoPE)
RoPE Dimensions	64	64	64	64

Table 2: The configuration settings of the Polyglot-Ko model.

3. 한국어 지시학습: 지시학습 데이터셋

지시문	오류 중 하나를 밝히세요.
입력	트위터, 인스타그램, 텔레그램
출력	트위터입니다.

지시문, 입력, 출력 형태의 데이터셋

지시문	세 가지 기본 색은 무엇인가요?
출력	기본 색은 빨강, 파랑, 노랑입니다.

지시문, 출력 형태의 데이터셋

학습 데이터셋

- KoAlpaca v1.0 데이터셋
- Self-Instruct 데이터셋을 기반으로 제작
- 지시문, 입력 부분은 번역기를 통해 번역
- 번역된 지시문과 출력으로 Open-AI의 "gpt-3.5-turbo"를 사용하여 출력 생성

3. 한국어 지시학습: 지시학습 데이터셋

지시문: 문법적으로 올바른 문장을 구성합니다.

Before

갈색 여우가 빨리 뛰었습니다.

After

여우가 갈색 뛰었습니다 빨리

학습 데이터셋 수정

- 문법 교정, 오류 고정을 위한 지시문이 번역기를 거치며 오류가 고정된 상태로 번역되는 문제가 존재
- 문법적으로 올바르지 않은 데이터셋으로 변경
- 전체 데이터셋 49,620개에서 " 오류, 교정, 문법" 을 키워드로 검색하여 82개의 데이터셋을 수정

3. 한국어 지시학습: 실험 세팅

프롬프트(prompt)는 입력 텍스트나 명령어를 의미하며 사용자가 컴퓨터와 상호작용 할 때 명령어나 질문 등을 포함하는 텍스트를 입력하는데 사용된다. 데이터셋의 입력 유무에 따라서 2가지 프롬프트를 사용하였다.

- `prompt_input`: 아래는 작업을 설명하는 명령어와 추가적 맥락을 제공하는 입력이 짝을 이루는 예제입니다. 요청을 적절히 완료하는 응답을 작성하세요. 명령어: “지시문” 입력: “입력” 응답: “출력”
- `prompt_no_input`: 아래는 작업을 설명하는 명령어입니다. 명령어에 따른 요청을 적절히 완료하는 응답을 작성하세요. 명령어: “지시문” 응답: “출력”

Prompt 예시

학습에는 NVIDIA A100-PCIE-40GB 8장을 사용하였다. 옵티마이저는 AdamW를 사용하였으며, deepspeed stage 3를 적용하여 학습하였다.

이 외의 하이퍼 파라미터 구성은 다음 표4와 같다.

하이퍼 파라미터	값
Epochs	3
Learning Rate	5e-5, 7e-6
Total Batch Size	32
Cosine Warmup Ratio	0.06
Gradient Clipping	1

학습에 사용한 하이퍼 파라미터

3. 한국어 지시학습: 실험 결과

평가를 위해 AI-hub-요약문 및 레포트 생성 데이터, nsmc 데이터셋을 사용하였다. AI-hub-요약문 및 레포트 생성 데이터는 생성된 결과와 정답과의 Rouge-score를 측정하였고, nsmc 데이터는 incontext-oneshot을 사용하여 Accuracy를 측정하였다. 요약문 및 레포트 생성 데이터의 프롬프트는 “주어진 기사를 요약하세요 기사: (문서)”를 사용하였고, nsmc 데이터의 프롬프트는 “문장의 긍정, 부정을 분류하세요. (문장) 정답: (감정)”의 형태를 사용하였다.

		AI-hub-요약문		NSMC
Model	LR	Rouge-1	Rouge-L	Accuracy
Polyglot-5.8b	5e-5	0.10	0.08	38.1
	7e-6	0.16	0.12	38.1
Polyglot-12.8b	5e-5	0.10	0.08	54.8
	7e-6	0.15	0.12	64.1

표 4: 성능 표

실험 결과 및 분석

- AI-hub-요약문 및 NSMC 데이터셋을 사용하여 지시학습 모델 성능 평가
- NSMC 데이터는 in-context-oneshot을 사용하여 Accuracy 측정
- “요약문 및 레포트 생성”의 경우 언어모델의 크기가 커짐에 따라 성능도 증가
- 언어모델이 커질수록 학습률에 따른 성능 차이가 있는 것으로 보이며 본 실험에서는 7e-6 상태일때, 더 좋은 성능을 확인 가능

감사합니다.