

SOLAR 10.7B: Scaling Large Language Models with Simple yet Effective Depth Up-Scaling

Background



✓ LLM의 도입은 NLP분야에 있어 엄청난 혁신을 가져다 줬으며 자연어에 대한 이해와 상호작용을 가능하게 함

✓ 따라서 LLM의 성능을 높이기 위해 더 많은 파라미터, MOE같은 Up-Scaling 방법 등이 제안되었음

✓ 그러나 MOE방식은 Train, Inference Framework를 변경해야 한다는 치명적인 단점 존재

✓ 이에, Depth up-scaling 방식을 통해 효과적으로 스케일링을 하는 것이 논문의 목표



Background

T	Model	Average	ARC	HellaSwag	MMLU	TruthfulQA
	kyujinpy/Sakura-SOLAR-Instruct	74.4	70.99	88.42	66.33	71.79
	Weyaxi/SauerkrautLM-UNA-SOLAR-Instruct-test	74.26	70.9	88.3	66.15	71.8
	Weyaxi/SauerkrautLM-UNA-SOLAR-Instruct	74.26	70.9	88.3	66.15	71.8
	VAG0solutions/SauerkrautLM-SOLAR-Instruct	74.21	70.82	88.63	66.2	71.95
	upstage/SOLAR-10.7B-Instruct-v1.0	74.2	71.08	88.16	66.21	71.43
	fblgit/UNA-SOLAR-10.7B-Instruct-v1.0	74.2	70.56	88.18	66.08	72.05
	fblgit/UNA-SOLAR-10.7B-Instruct-v1.0	74.07	70.73	88.32	66.1	72.52
	Yhyu13/LMCocktail-10.7B-v1	74.06	70.65	88.13	66.21	71.03
	yhyu13/LMCocktail-10.7B-v1	74.06	70.65	88.13	66.21	71.03
	rishiraj/meow	73.94	70.48	88.08	66.25	70.49
	fblgit/LUNA-SOLARxautLM-Instruct	73.79	71.16	88.28	66.11	73.37

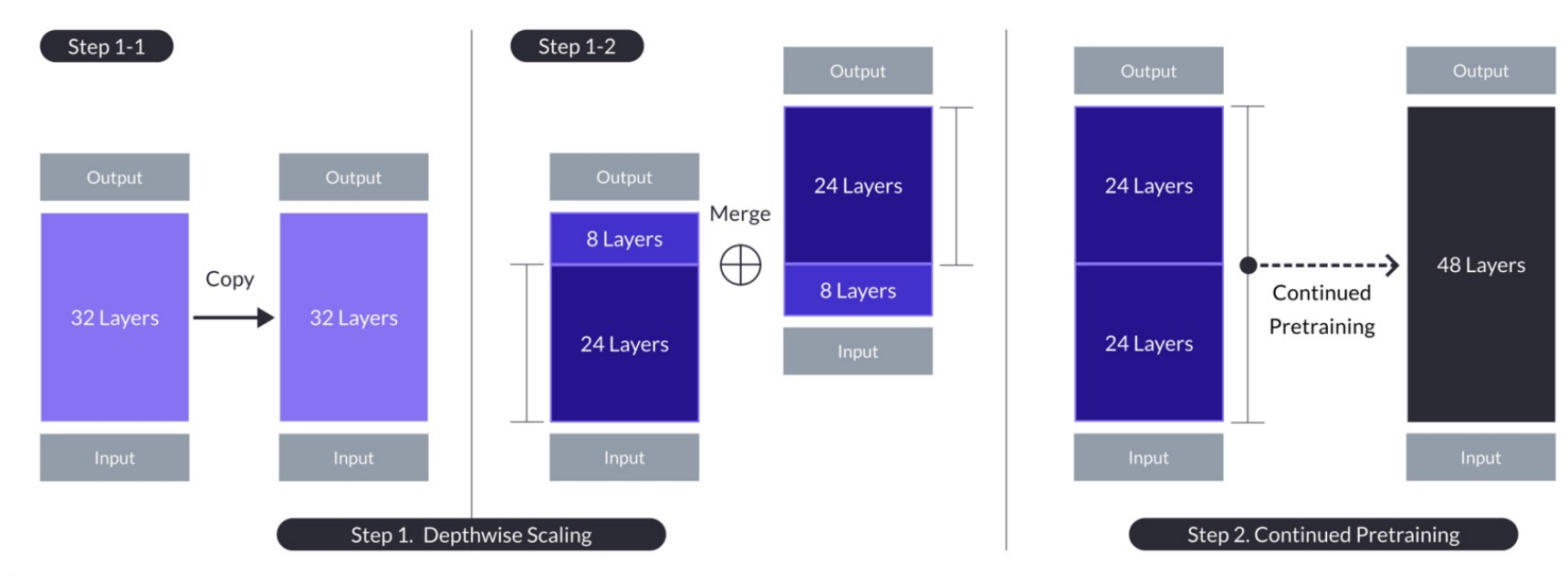
	davidkim205/Rhea-72b-v0.5	81.22	79.78	91.15
	MTSAIR/MultiVerse_70B	81	78.67	89.77
	MTSAIR/MultiVerse_70B	80.98	78.58	89.74
	abacusai/Smaug-72B-v0.1	80.48	76.02	89.27
	ibivibiv/alpaca-dragon-72b-v1	79.3	73.89	88.16
	mistralai/Mixtral-8x22B-Instruct-v0.1	79.15	72.7	89.08
	MaziyarPanahi/Llama-3-70B-Instruct-DPO-v0.2	78.96	72.53	86.22
	MaziyarPanahi/Llama-3-70B-Instruct-DPO-v0.4	78.89	72.61	86.03
	MaziyarPanahi/Llama-3-70B-Instruct-DPO-v0.3	78.74	72.35	86
	mmnga/Llama-3-70B-japanese-suzume-vector-v0.1	78.6	72.35	85.81
	moreh/MoMo-72B-lora-1.8.7-DPO	78.55	70.82	85.96
	tenyx/Llama3-TenyxChat-70B	78.4	72.1	86.21

Approach

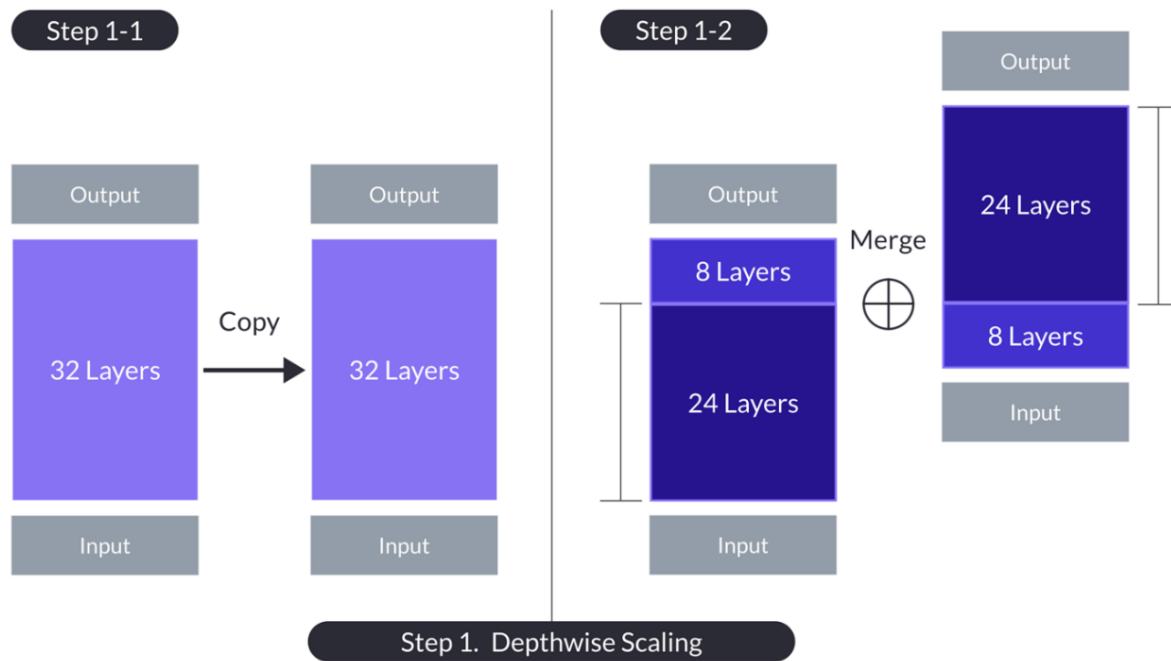
	DUS (Depth Up-Scaling)	MoE (Mixture of Experts)
개념	기존 모델의 층(layer)을 복제하여 모델의 깊이를 증가시키는 방식	여러 전문가(expert) 모듈을 결합하여 모델의 용량을 확장하는 방식
구현의 복잡성	상대적으로 간단	복잡한 구조와 추가적인 모듈 필요
훈련 및 추론 프레임워크	기존 훈련 및 추론 프레임워크와 호환	별도의 훈련 프레임워크 및 특수한 추론 메커니즘 필요
모델 호환성	모든 트랜스포머 아키텍처와 호환	특정 아키텍처에 맞게 조정 필요
효율성	높음 (기존 가중치 재사용)	중간 (전문가 모듈에 따라 다름)
확장성	층을 늘리는 방식으로 확장	전문가 모듈을 추가하여 확장

출처: <https://dajeblog.co.kr/%EB%85%BC%EB%AC%B8-%EB%A6%AC%EB%B7%B0-solar-10-7b-scaling-large-language-models-with-simple-yet-effective-depth-up-scaling/>

Architecture



Architecture - Base Model, Depth up Scaling



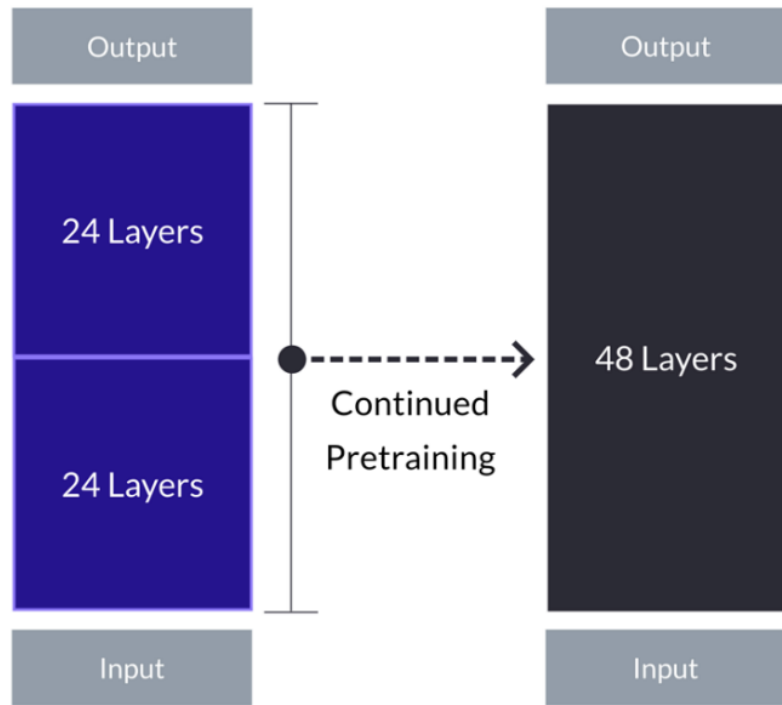
- Base Model

- ✓ 32 Layer Llama2 Architecture 사용
- ✓ 이를 SOTA 성능을 가진 Mistral7B의 가중치로 초기화

- Depth up Scaling

1. Base Model을 복제하여 두 개의 모델 생성
2. Base Model은 처음 8개의 Layer 제거, 다음 모델은 마지막 8개의 Layer 제거
3. 두 개의 모델을 연결함으로써 총 42개 Layer 모델 완성

Architecture - Continued Pretraining



Step 2. Continued Pretraining

- ✓ 레이어를 제거하였기 때문에 초기의 성능을 매우 떨어짐
- ✓ 그렇기 위해 Pretraining을 반복하게 되며 이에 성능이 빠르게 회복되는 것을 확인

**** 왜 굳이 레이어를 제거하지, 강 연결시키면 되는거 아닌가?**

중간에 레이어를 제거함으로써 레이어간의 거리, 즉 정보의 차이를 줄이는 거임!!

저자들은 이러한 방법이 본인들 모델의 성능에 엄청난 도움을 주었음을 확신함

Architecture - instruction tuning, alignment tuning

instruction tuning

- ✓ 모델에게 Instruction을 줌으로써 모델의 Output을 제한하고 Specific한 Task 수행 가능
- ✓ QA 포맷으로 된 Instruction Data 사용
- ✓ 대부분 오픈소스 데이터셋을 사용했지만 수학능력의 상승을 위해 Math QA Dataset도 사용

instruction:

Give three tips for staying healthy.

output:

1. Eat a balanced and nutritious diet: Make sure your meals are inclusive of a variety of fruits and vegetables, lean protein, whole grains, and healthy fats. This helps to provide your body with the essential nutrients to function at its best and can help prevent chronic diseases. 2. Engage in regular physical activity: Exercise is crucial for maintaining strong bones, muscles, and cardiovascular health. Aim for at least 150 minutes of moderate aerobic exercise or 75 minutes of vigorous exercise each week. 3. Get enough sleep: Getting enough quality sleep is crucial for physical and mental well-being. It helps to regulate mood, improve cognitive function, and supports healthy growth and immune function. Aim for 7-9 hours of sleep each night.

alignment tuning

- ✓ 선호와 align 되도록 RLHF의 기법 중 하나인 Direct Preference Optimization 학습 수행
- ✓ 이는 기존의 RLHF 기법보다 계산 비용 측면에서 이점이 있으며 성능 또한 뛰어남
- ✓ 마찬가지로 성능을 위해 Math-Instruct Dataset 사용

system:

You are an AI assistant. You will be given a task. You must generate a detailed and long answer.

question:

Generate an approximately fifteen-word sentence that describes all this data: Midsummer House eatType restaurant; Midsummer House food Chinese; Midsummer House priceRange moderate; Midsummer House customer rating 3 out of 5; Midsummer House near All Bar One

chosen:

Midsummer House is a moderately priced Chinese restaurant with a 3/5 customer rating, located near All Bar One.

rejected:

Sure! Here's a sentence that describes all the data you provided: "Midsummer House is a moderately priced Chinese restaurant with a customer rating of 3 out of 5, located near All Bar One, offering a variety of delicious dishes."

Experiments

Main Results

Model	Size	Type	H6 (Avg.)	ARC	HellaSwag	MMLU	TruthfulQA	Winogrande	GSM8K
SOLAR 10.7B-Instruct	~ 11B	Alignment-tuned	74.20	71.08	88.16	66.21	71.43	83.58	64.75
Qwen 72B	~ 72B	Pretrained	73.60	65.19	85.94	77.37	60.19	82.48	70.43
Mixtral 8x7B-Instruct-v0.1	~ 47B	Instruction-tuned	72.62	70.22	87.63	71.16	64.58	81.37	60.73
Yi 34B-200K	~ 34B	Pretrained	70.81	65.36	85.58	76.06	53.64	82.56	61.64
Yi 34B	~ 34B	Pretrained	69.42	64.59	85.69	76.35	56.23	83.03	50.64
Mixtral 8x7B-v0.1	~ 47B	Pretrained	68.42	66.04	86.49	71.82	46.78	81.93	57.47
Llama 2 70B	~ 70B	Pretrained	67.87	67.32	87.33	69.83	44.92	83.74	54.06
Falcon 180B	~ 180B	Pretrained	67.85	69.45	88.86	70.50	45.47	86.90	45.94
SOLAR 10.7B	~ 11B	Pretrained	66.04	61.95	84.60	65.48	45.04	83.66	55.50
Qwen 14B	~ 14B	Pretrained	65.86	58.28	83.99	67.70	49.43	76.80	58.98
Mistral 7B-Instruct-v0.2	~ 7B	Instruction-tuned	65.71	63.14	84.88	60.78	68.26	77.19	40.03
Yi 34B-Chat	~ 34B	Instruction-tuned	65.32	65.44	84.16	74.90	55.37	80.11	31.92
Mistral 7B	~ 7B	Pretrained	60.97	59.98	83.31	64.16	42.15	78.37	37.83

✓ 다른 Top Model들과의 6가지 Task 정확도 비교, SOLAR 10.7B-Instruct가 SOTA 달성

✓ MOE를 수행한 Mistral 7B, Qwen 14B에 비해 성능이 높은 것을 확인, 이는 DUS가 Up-scaling에서 효과적임을 증명

Experiments

1. Instruct tuning

Model	Alpaca-GPT4	OpenOrca	Synth. Math-Instruct	H6 (Avg.)	ARC	HellaSwag	MMLU	TruthfulQA	Winogrande	GSM8K
SFT v1	○	✗	✗	69.15	67.66	86.03	65.88	60.12	82.95	52.24
SFT v2	○	○	✗	69.21	65.36	85.39	65.93	58.47	82.79	57.32
SFT v3	○	○	○	70.03	65.87	85.55	65.31	57.93	81.37	64.14
SFT v4	○	✗	○	70.88	67.32	85.87	65.87	58.97	82.48	64.75
SFT v3 + v4	○	○	○	71.11	67.32	85.96	65.95	58.80	82.08	66.57

- ✓ Alpaca-GPT4 Dataset만을 사용한 SFT v1의 H6 점수는 69.15
- ✓ OpenOrca dataset을 추가한 SFT v2는 평균점수는 비슷하지만 GSM8K에서 급격히 증가함과 동시에 ARC 등의 점수가 하락, 이는 이 데이터셋이 모델의 결과에 영향을 미침을 알 수 있음
 - ✓ Synth 데이터셋을 추가한 SFT v3는 GSM8K 성능을 엄청나게 끌어올림, 흥미로운 점은 SFT v3 + v4가 모든 면에서 SFT v3의 점수를 뛰어넘는다는 것
 - ✓ SFT v3 + v4는 단순히 모델의 가중치를 평균 내어 병합한 것으로 성능이 가장 좋음을 확인

Experiments

2. Alignment tuning

Model	Ultrafeedback Clean	Synth. Math-Alignment	H6 (Avg.)	ARC	HellaSwag	MMLU	TruthfulQA	Winogrande	GSM8K
DPO v1	○	✗	73.06	71.42	88.49	66.14	72.04	81.45	58.83
DPO v2	○	○	73.42	71.50	88.28	65.97	71.71	82.79	60.27
DPO v1 + v2	○	○	73.21	71.33	88.36	65.92	72.65	82.79	58.23

- ✓ DPO v1 + v2 모델 역시 같은 방법으로 병합
- ✓ DPO V1의 경우 H6 73.06으로 SFT의 모든 모델들보다 성능이 향상되었지만 GSM8K만의 성능은 하락함을 확인
 - ✓ DPO v1 + v2의 경우 성능이 증가하지 않았으며 핵심은 Synth. Math-Alignment의 추가임을 알 수 있음

Conclusion

limitation

- ✓ 하드웨어의 한계로 인해 8개의 Layer를 제거했으나 이게 가장 효과적인 수치인지는 확인하지 못했음
- ✓ 훈련 시간의 증가와 계산 비용의 문제로 계속된 실험을 진행하지 못함
- ✓ 모든 딥러닝 모델과 마찬가지로 훈련 데이터의 Bias문제를 해결하지 못함
- ✓ 아주 특화된 Task를 수행하기 위해서는 더욱 더 특화된 Instruct data가 필요

Conclusion

- ✓ DUS 방식의 모델이 성능이 매우 뛰어남을 보여줬으며 이는 NLP Tasks에서 의미 있는 적용이 될 수 있음
- ✓ 10.7B개의 파라미터만으로도 다른 거대 언어 모델보다 성능이 뛰어남을 입증

감사합니다 😊