



LLM

2024. 9.

Industrial Data Science Lab

I. 생성형 AI : LLM

• Scaling Laws (Open AI, 2020)

- 컴퓨팅 리소스, 데이터, 모형 크기를 늘릴 수록 성능 개선
- 새로운 능력이 생겨남 (Emergent Abilities)

Scaling Laws for Neural Language Models			
Jared Kaplan *		Sam McCandlish*	
Johns Hopkins University, OpenAI		OpenAI	
jaredk@jhu.edu		sam@openai.com	
Tom Henighan	Tom B. Brown	Benjamin Chess	Rewon Child
OpenAI	OpenAI	OpenAI	OpenAI
henighan@openai.com	tom@openai.com	bchess@openai.com	rewon@openai.com
Scott Gray	Alec Radford	Jeffrey Wu	Dario Amodei
OpenAI	OpenAI	OpenAI	OpenAI
scott@openai.com	alec@openai.com	jeffwu@openai.com	damodei@openai.com

Abstract

We study empirical scaling laws for language model performance on the cross-entropy loss. The loss scales as a power-law with model size, dataset size, and the amount of compute used for training, with some trends spanning more than seven orders of magnitude. Other architectural details such as network width or depth have minimal effects within a wide range. Simple equations govern the dependence of overfitting on model/dataset size and the dependence of training speed on model size. These relationships allow us to determine the optimal allocation of a fixed compute budget. Larger models are significantly more sample-efficient, such that optimally compute-efficient training involves training very large models on a relatively modest amount of data and stopping significantly before convergence.

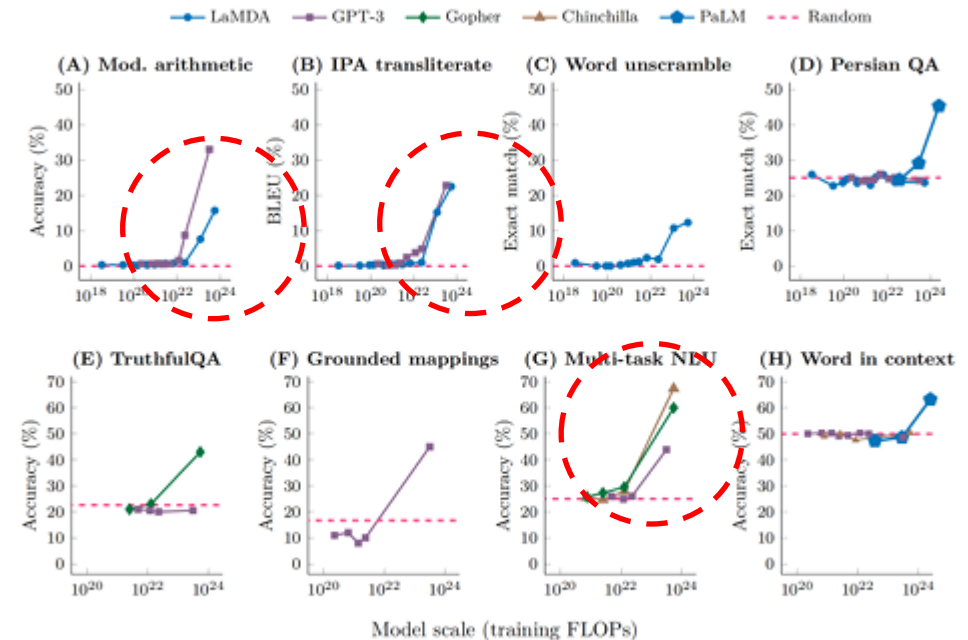
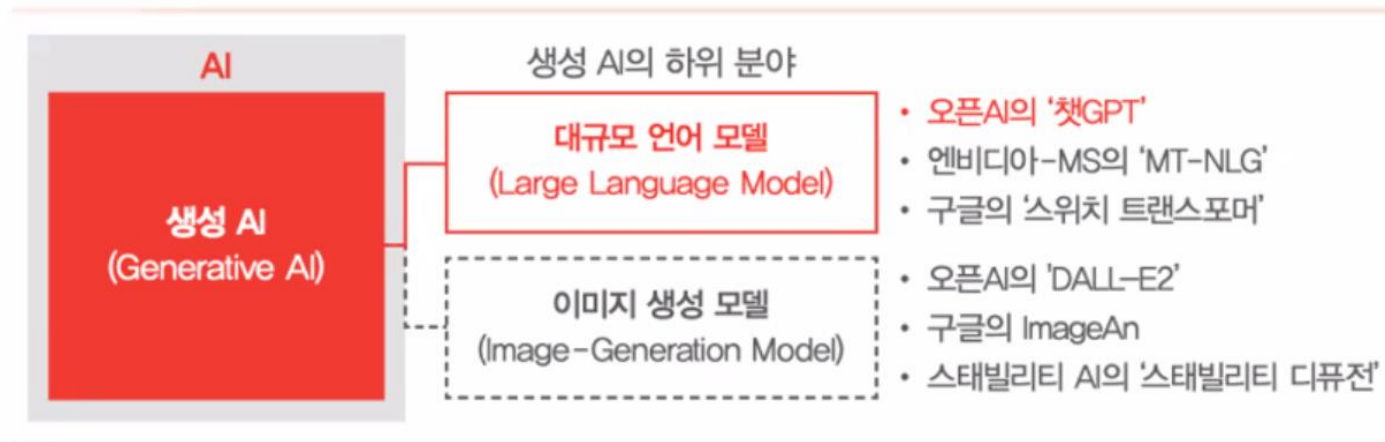


Figure 1: Emergent abilities of large language models. Model families display *sharp* and *unpredictable* increases in performance at specific tasks as scale increases. Source: Fig. 2 from [33].

• 생성형 AI

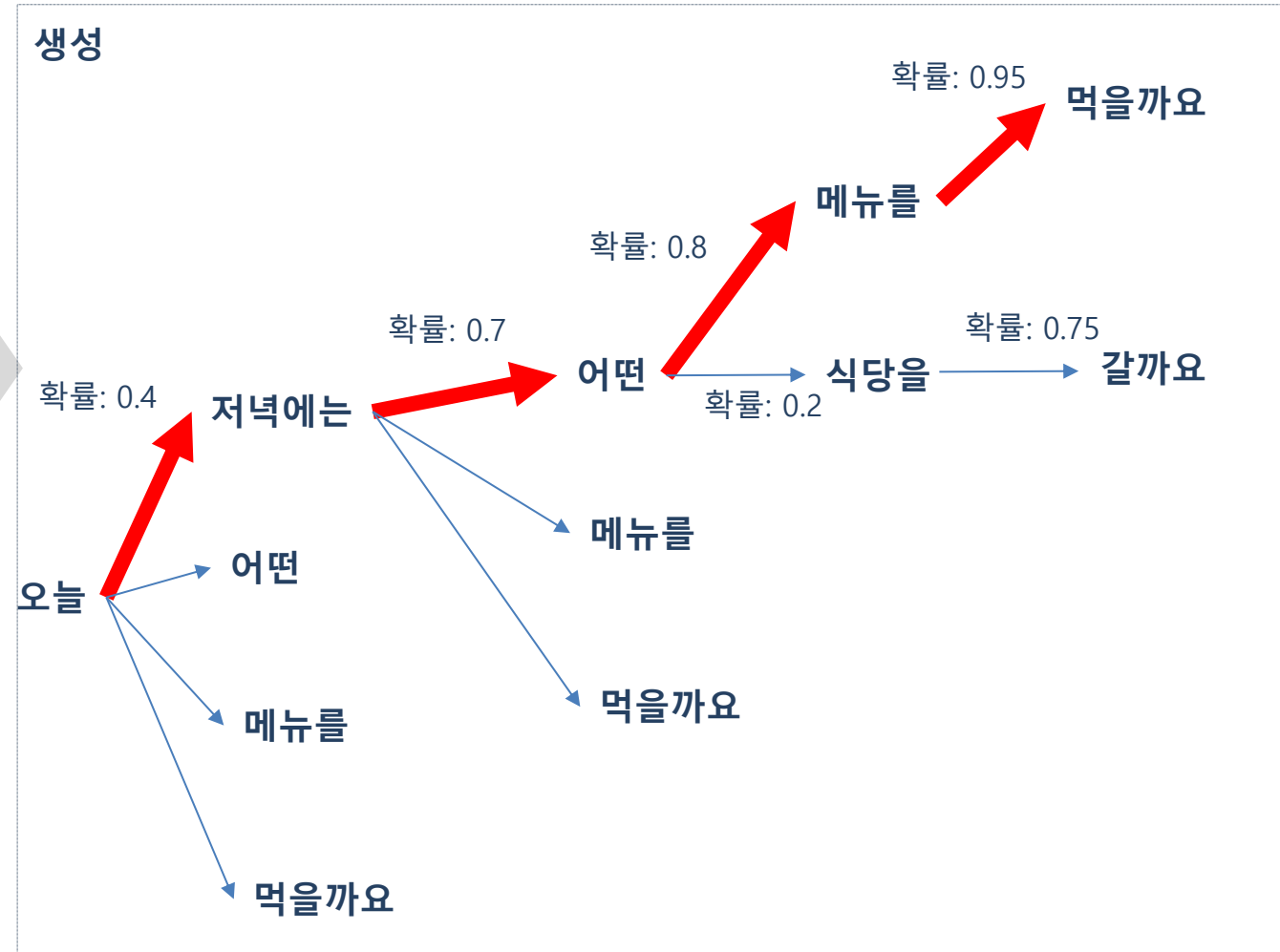
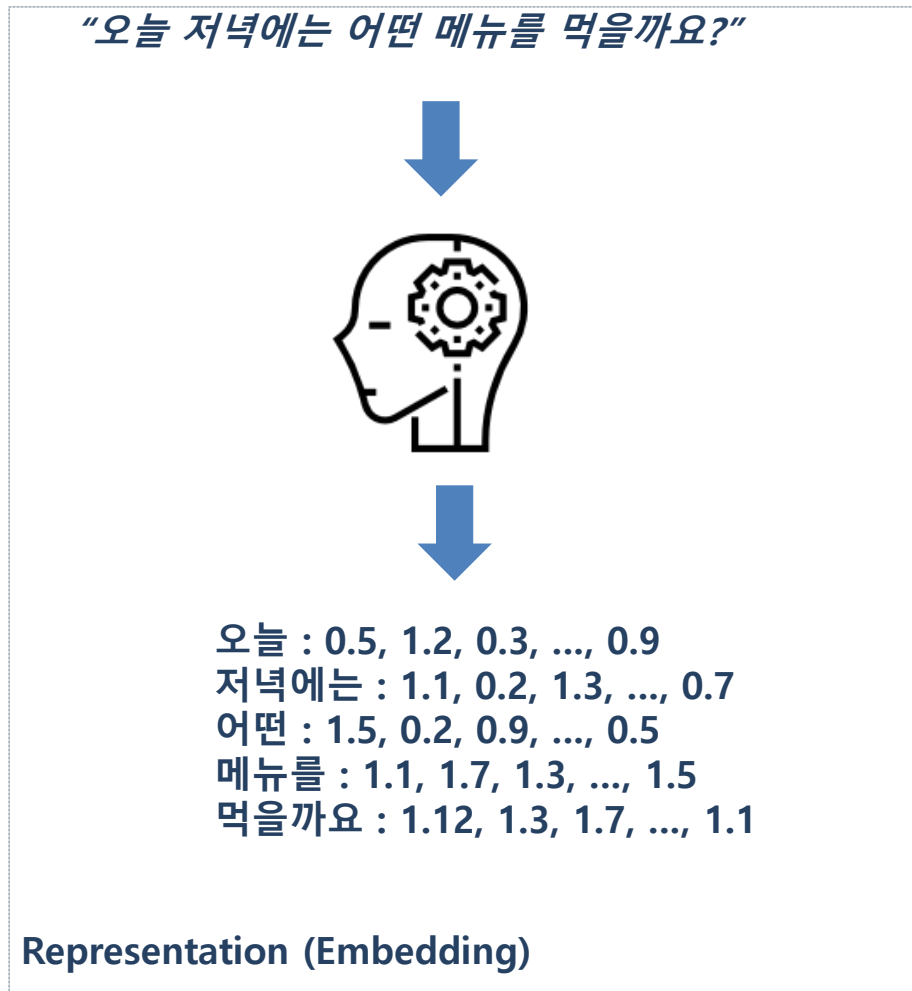
- 대형 언어모델이나 이미지 생성모델 활용, User가 원하는 것을 생성하는 AI 분야
- LLM: 테라바이트 단위의 대용량 코퍼스로 모형을 학습, 문장을 생성하는 모형
- IGM(Image Generation Model): 텍스트를 입력하면 그에 대한 이미지를 생성, DALL-E2, 미드저니 등



※ 자료: 삼일PwC경영연구원

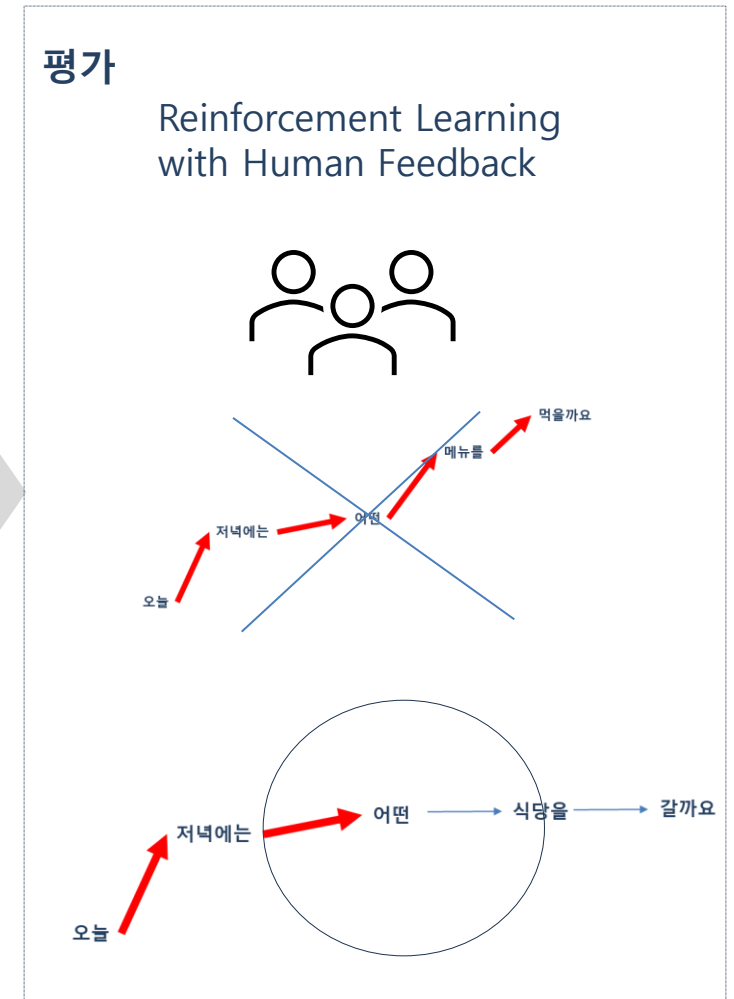
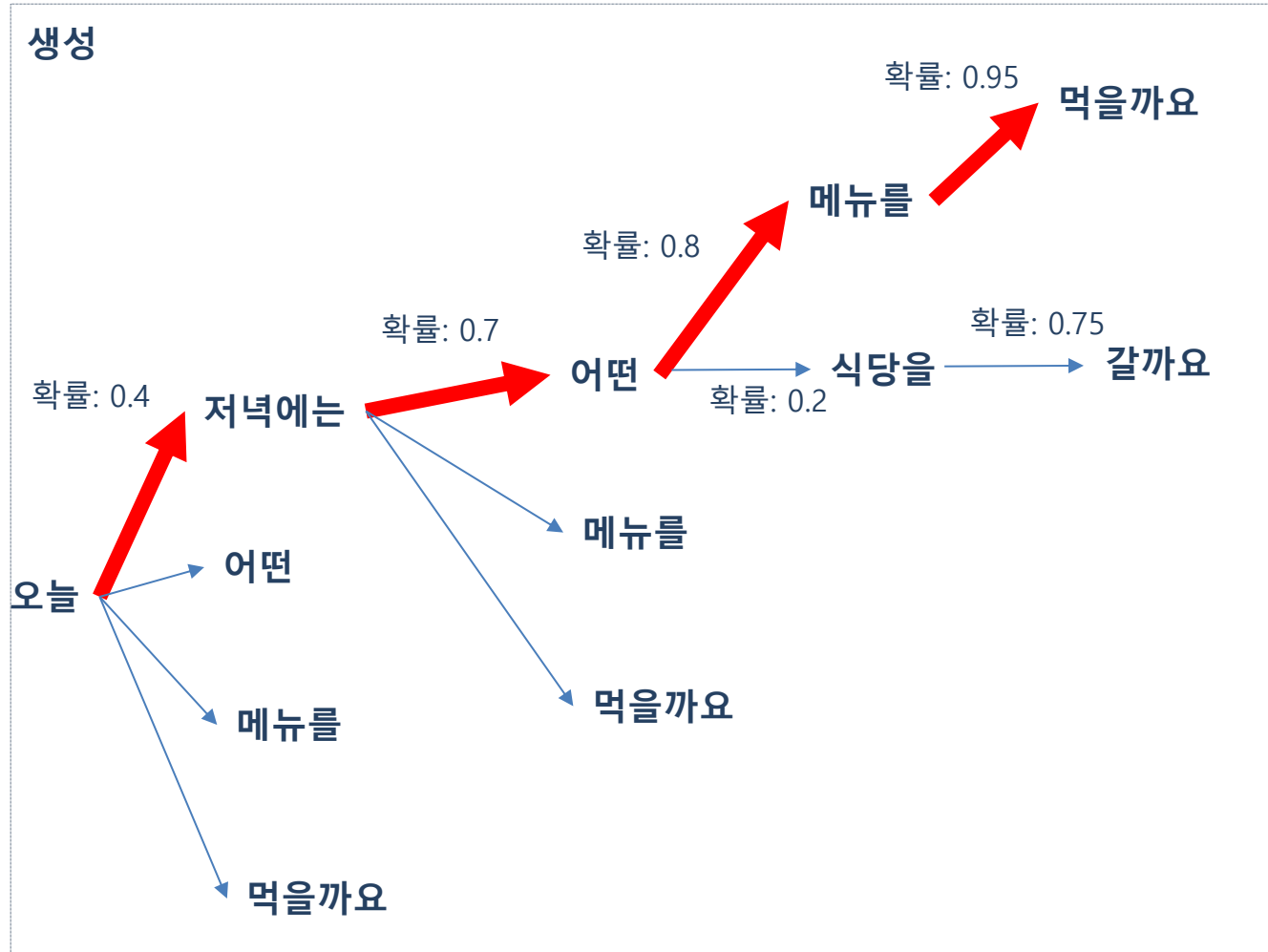
I. 생성형 AI : LLM

• 간단하게 이해하는 LLM



I. 생성형 AI : LLM

• 간단하게 이해하는 LLM



• 챗GPT

- GPT3는 GPT2에 비해 10배 이상의 데이터(570GB), 100배 이상 모형 크기(175B)
- GPT4는 1700B의 파라미터(estimated)
- Few-shot 학습을 통한 In-Context Learning 향상
 - 프롬프트의 내용만으로 Task 수행
 - 프롬프트 내 맥락을 이해하고 답변
- GPT4o 출시

 You

나에게 송년회 건배사를 3개만 알려줘, 건배사가 감동적이었으면 좋겠어

 ChatGPT

물론, 아래는 감동적인 송년회 건배사의 예시 중 3가지입니다. 이 중 하나를 선택하여 송년회에서 사용하실 수 있습니다.

1. 기회와 변화를 의미있게 채우며:

"우리는 또 하나의 해를 마무리하고 새로운 시작을 앞두고 있습니다. 이번 해도 여러 어려움과 도전에 직면했지만, 그런 순간들에서 우리는 더 강해지고 성장했습니다. 이제 하는 일과 목표는 새로운 해를 기대하게 만듭니다. 모두가 함께 협력하고 서로를 지원하면, 무엇이든 이룰 수 있습니다. 새해에는 우리 모두가 행복하고 건강하게 지내길 기원하며, 함께 하는 모든 순간을 소중히 여기고 감사하게 생각합니다. 건배!"

2. 우정과 고마움 표현:

"이 자리에 모여 함께 시간을 보낸 모든 분들에게 감사의 말씀을 전하고 싶습니다. 우리는 서로를 향한 그 무한한 우정을 실제로 느낄 수 있었던 해였습니다. 다음 해에도 함께 소중한 순간을 나누고, 어려움을 함께 극복하며, 더 많은 기억을 만들어 나갈 것입니다. 모두에게 행복하고 풍요로운 새해를 기원합니다. 건배!"

3. 빛나는 미래를 기대하며:

"저희는 이 송년회에서 과거를 되돌아보고, 현재를 즐기며 미래를 기대합니다. 우리는 계속해서 발전하고 성장하며, 새로운 도약을 위해 노력할 것입니다. 이곳에 모여주신 모든 분들에게 감사의 인사를 전하며, 다음 해에는 더 큰 성취와 행복을 경험하길 바랍니다. 새해에는 우리의 꿈을 이루고 빛나는 미래를 함께 만들어 나가기를 기대합니다. 건배!"

II. 생성형 AI 활용

- 프롬프트 구조화

- Delimiter와 프롬프트 영역 지정

- Delimiter: 특정 토큰으로 프롬프트에서 영역을 지정할 때 활용
 - ###, >>>, ==, <태그> 등으로 표현, 각 영역이 의도에 맞게 챗GPT에 전달되도록 함

- 활용 예시 프롬프트

대화 예시와 결과 예시를 참고해서, 다음 <<<대화>>>에 나오는 각 문장의 감성을 '긍정' 이나 '부정' 으로 분류하시오.

###

대화 예시

[행인1]: 오늘 날씨가 좋네요

[행인2]: 좋은 뭐가 좋아요

###

결과 예시

긍정

부정

###

<<<

[행인1] 비가 오니 차분합니다.

[행인2] 비오는 날 분위기 좋아요

>>>

II. 생성형 AI 활용

- 프롬프트 구조화

- XML 태그 스타일의 영역 지정

- XML태그?

- XML은 데이터의 구조를 기술하는 Mark up 언어
 - 데이터의 각 항목에 대해 의미하는 태그로 마크업
 - 예: <키워드> 챗, 생성AI </키워드>

II. 생성형 AI 활용

- 프롬프트 구조화

- 활용 예시 프롬프트

대화 예시와 결과 예시를 참고해서, 다음 <<<대화>>>에 나오는 각 문장의 감성을 '긍정' 이나 '부정' 으로 분류하시오.

<범주>

긍정

부정

</범주>

<대화예시>

[행인1]: 오늘 날씨가 좋네요

[행인2]: 좋은 뭐가 좋아요

</대화예시>

<결과예시>

긍정

부정

</결과예시>

<대화>

[행인1] 비가 오니 짜증납니다.

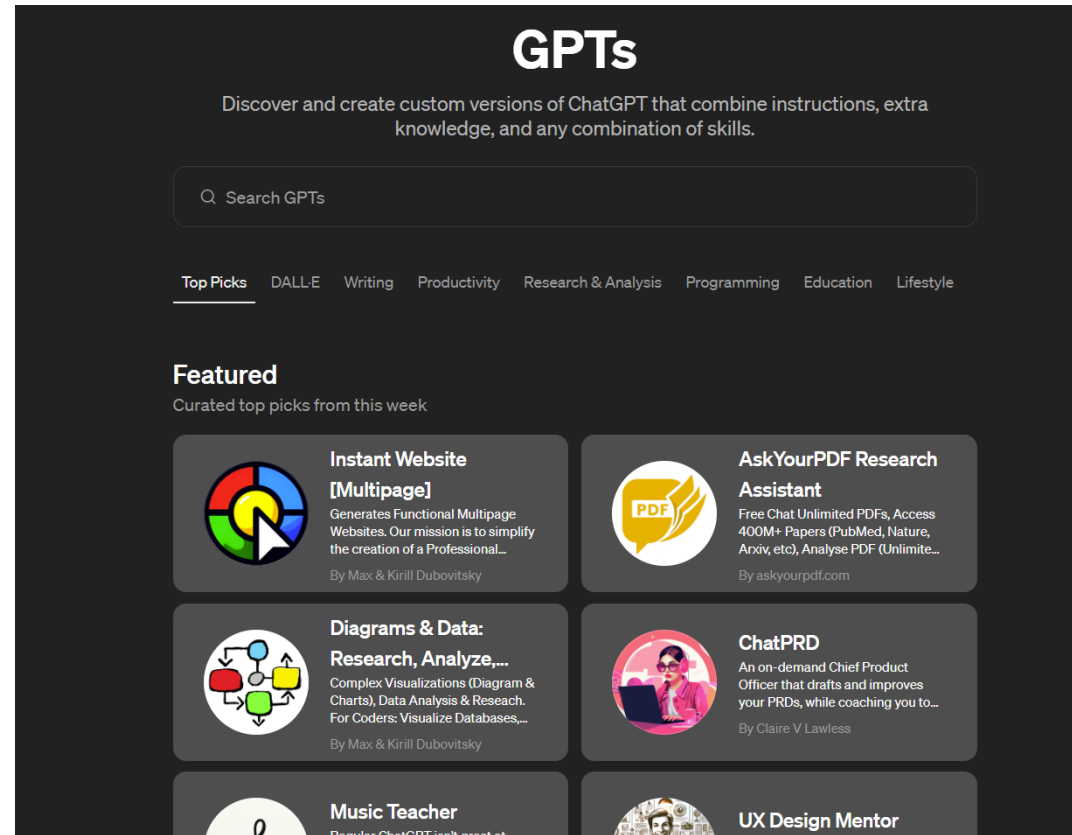
[행인2] 비오는 날 분위기 좋아요

</대화>

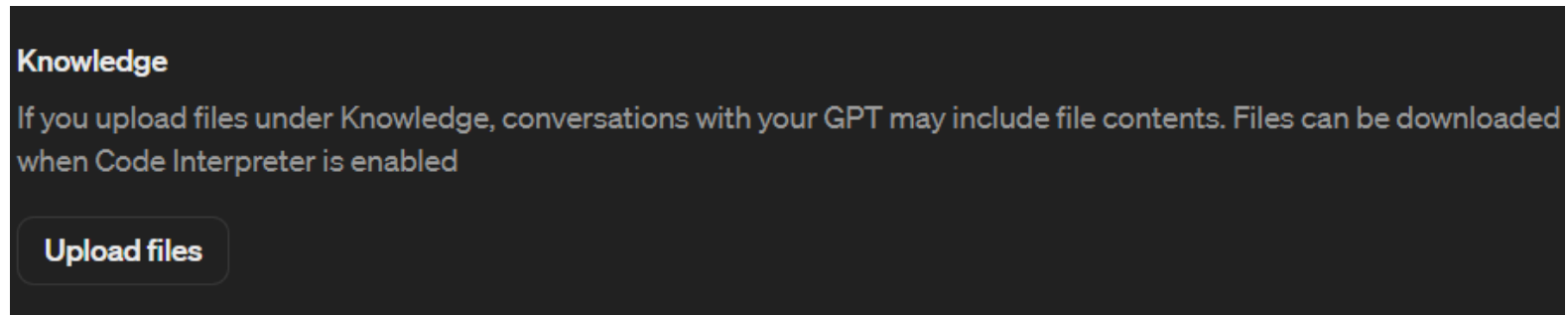
II. 생성형 AI 활용

- 챗GPT+커스텀GPTs

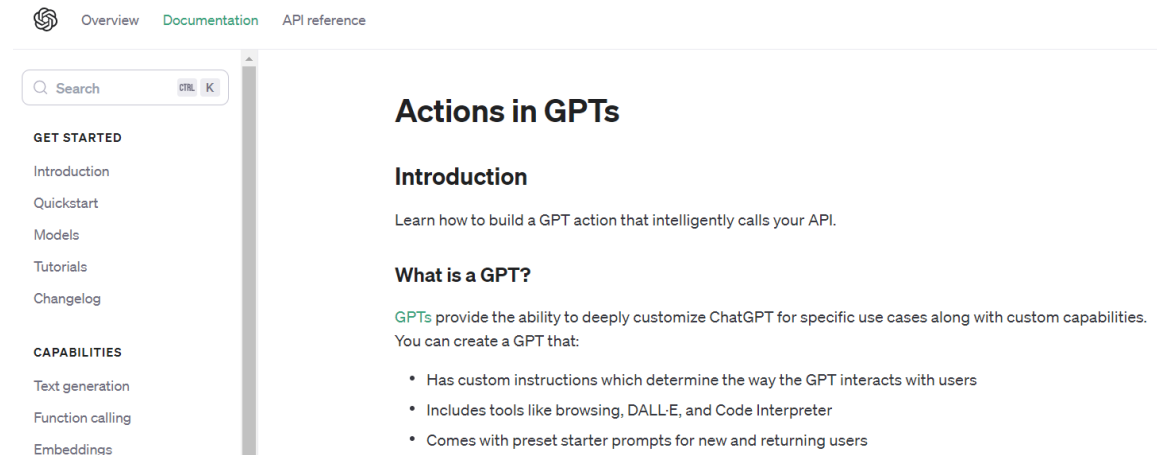
- GPTs: Customize된 챗GPT를 의미
- instruction: 제약사항, 가이드라인, 설명, 개인화 등을 확인
- GPT의 앱마켓



- 챗GPT+커스텀GPTs
 - Knowledge에 파일 업로드




- 외부 API의 활용



II. 생성형 AI 활용

- LMM(Large Multimodal Model)로의 발전

You



이 그림에서 이상한 부분을 패널별로 설명해주세요

ChatGPT

이 그림에는 세 개의 패널이 있고 각각 다른 장면을 보여주고 있습니다:

- 첫 번째 패널은 휴대폰과 같은 전자 장치가 나타나는데, 휴대폰의 충전 포트에 VGA 케이블이 연결되어 있는 것으로 보입니다. 일반적으로 VGA 케이블은 모니터와 같은 디스플레이 장치에 사용되며 휴대폰 충전에는 사용되지 않습니다.
- 두 번째 패널은 일본어로 된 패키지가 보이는데, 이것이 무엇인지는 명확하지 않습니다. 다만 패키지에는 전자기기나 기술 제품과 관련된 것처럼 보입니다.
- 세 번째 패널은 온도계를 손에 들고 있는 장면입니다. 온도계는 보통의 온도 측정 기기처럼 보이며, 다른 두 패널과의 직접적인 연관성은 보이지 않습니다.

이 세 장면 모두 상이한 맥락을 가지고 있어서, 전체적인 이야기나 상황을 이해하기 어렵습니다. 그리고 VGA 케이블이 휴대폰에 연결된 것은 기술적으로 말이 되지 않는 부분으로, 이 그림의 이상한 점 중 하나로 볼 수 있습니다.


- 프롬프트 엔지니어링의 전망

- Multi-modal AI로의 확장 (GPT4)
- 이미지 속 대상을 인식, 대상 간 관계를 파악하고 추론, 적절한 단어를 통해 답을 생성

You

위와 첫 패널과 유사한 다른 그림 만들어주세요

ChatGPT



Here is an illustration based on the description of the first panel you provided.



Industrial Data Science Lab

Contact:

won.sang.l@gmail.com

<https://sites.google.com/view/idslab>