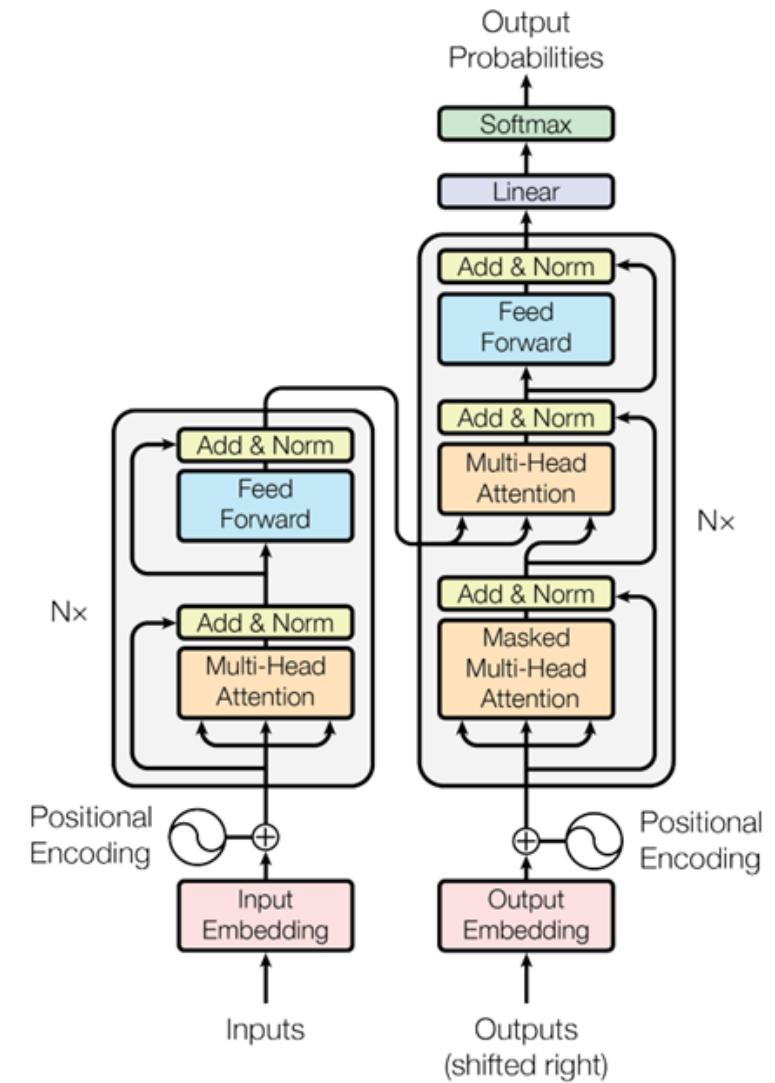


강원대학교
AI 소프트웨어학과

인공지능
- Chatgpt -

01 Chat-GPT

- 1950 : Rule-based Data Processing
- 1980 : Machine Learning
- 1990~2000 : Neural Network(RNN, LSTM)
- 2017~2018 : Transformer
- 2018~Now : GPT



Attention is all you need

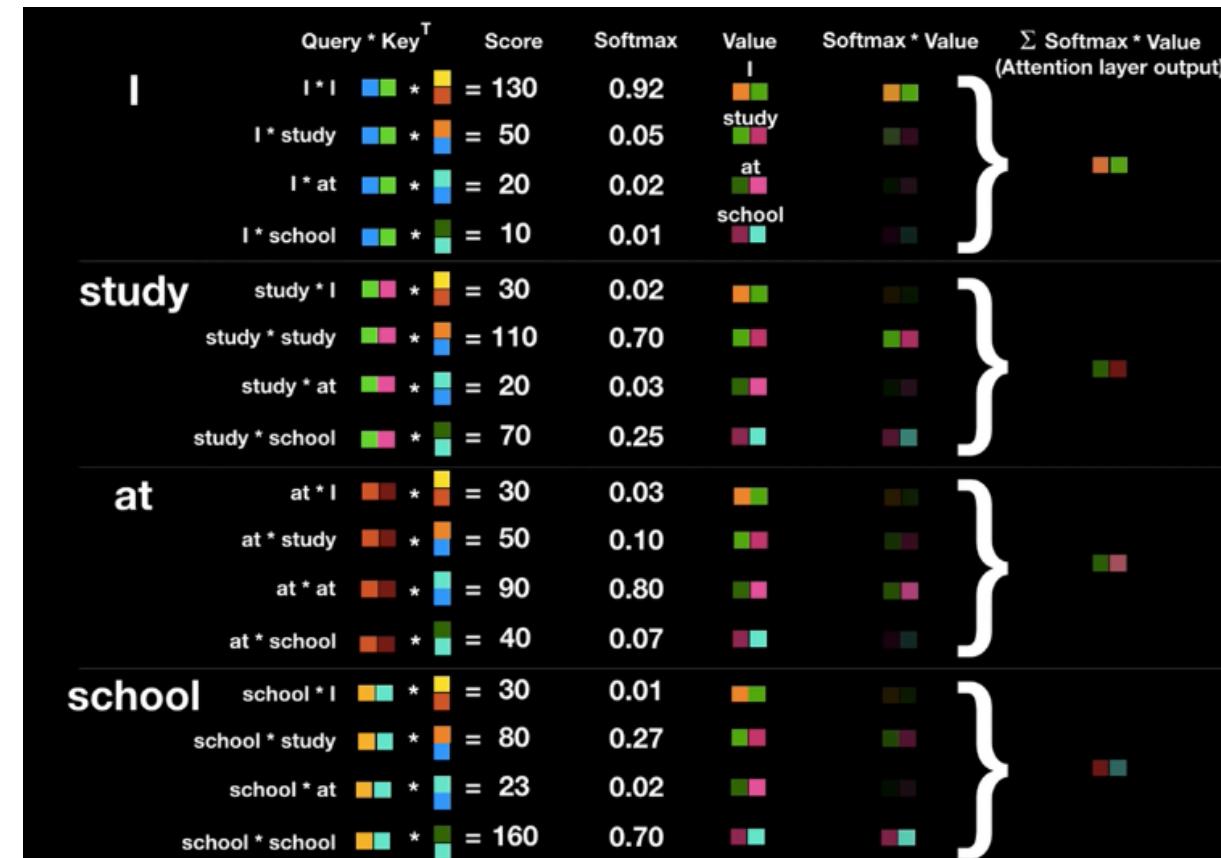
01 Chat-GPT

1990~2000 : Neural Network(RNN, LSTM)

- 자연어 처리 모델로 자연어에 대해 순차적으로 처리함
- 고정된 순서에 의해 제약을 받음
- 긴 문장을 이해할 수 없음

2018 : Transformer, BERT

- 자연어의 입력을 병렬로 처리하고 각 단어에 대한 연관 관계를 파악
- 문자들의 순서에 의해 제약 받지 않고 문자 간의 관계를 파악함
→ 양방향성 문맥을 파악하는데 유용함
- 긴 문장을 이해할 수 있음



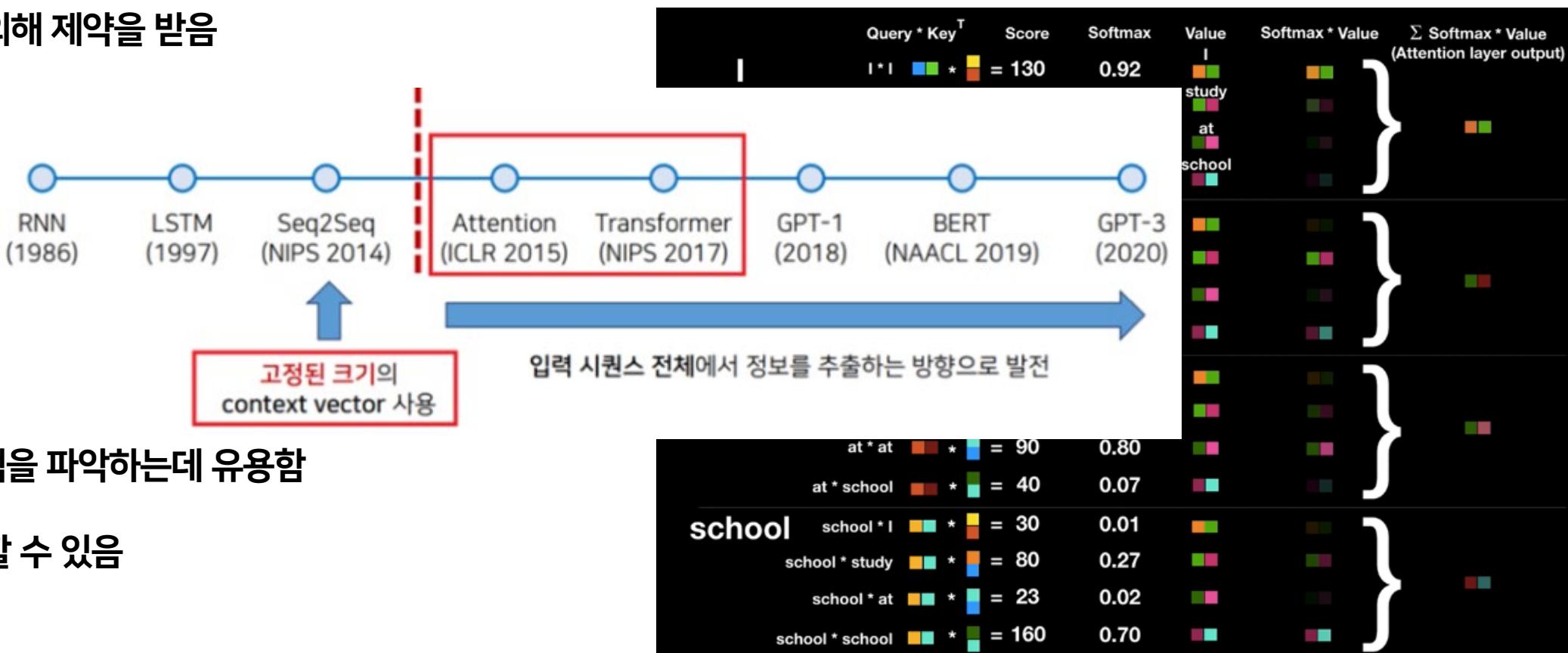
01 Chat-GPT

1990~2000 : Neural Network(RNN, LSTM)

- 자연어 처리 모델로 자연어에 대해 순차적으로 처리함
- 고정된 순서에 의해 제약을 받음
- 긴 문장을 이해

2018 : Transformer

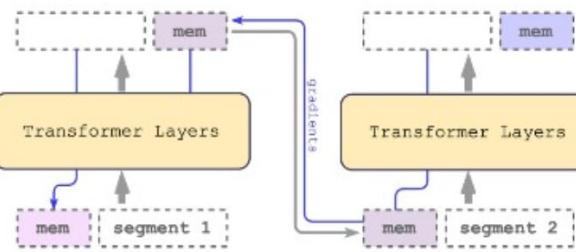
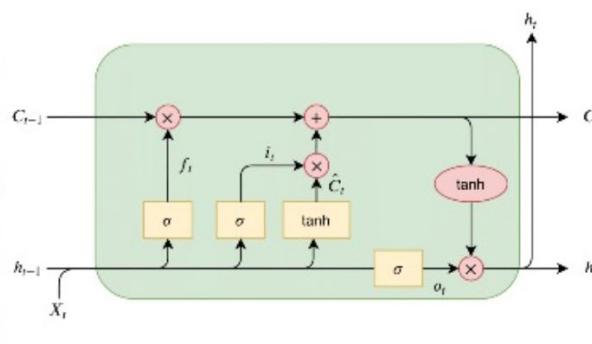
- 자연어의 입력
파악
- 문자들의 순서
→ 양방향성 문맥을 파악하는데 유용함
- 긴 문장을 이해할 수 있음



- The cat is sit on the mat"라는 문장에 대해 학습된 경우 "The"라는 단어를 처리하는 것으로 시작하여 "cat", "is", "sitting", "on", "the", 그리고 "mat". RNN의 숨겨진 상태는 현재 단어와 이전 숨겨진 상태를 기반으로 각 단계에서 업데이트됨
- I ate the apple because it was delicious. 라는 문장에서 전통적인 언어 모델은 중요한 컨텍스트를 제공하는 " ate "라는 단어에서 몇 단어 떨어져 있기 때문에 " delicious "라는 단어를 올바르게 예측하는 데 어려움
 - 누군가가 말하는 것을 듣거나 문장을 소리내어 읽을 때 단어를 한 번에 하나씩 순차적으로 처리함
- "The cat is sit on the mat"라는 동일한 문장에 대해 교육을 받은 경우 먼저 전체 문장을 고차원 벡터 공간에 삽입함
- Transformer 모델은 각 단어의 중요성을 효과적으로 평가하고, "ate"와 "delicious" 사이의 관계를 포착하여 정확한 예측이 가능
 - 긴 단락이나 기사를 읽을 때 텍스트의 다른 부분에 동시에 주의를 기울여 그 의미를 이해하는 능력을 사용할 수 있음

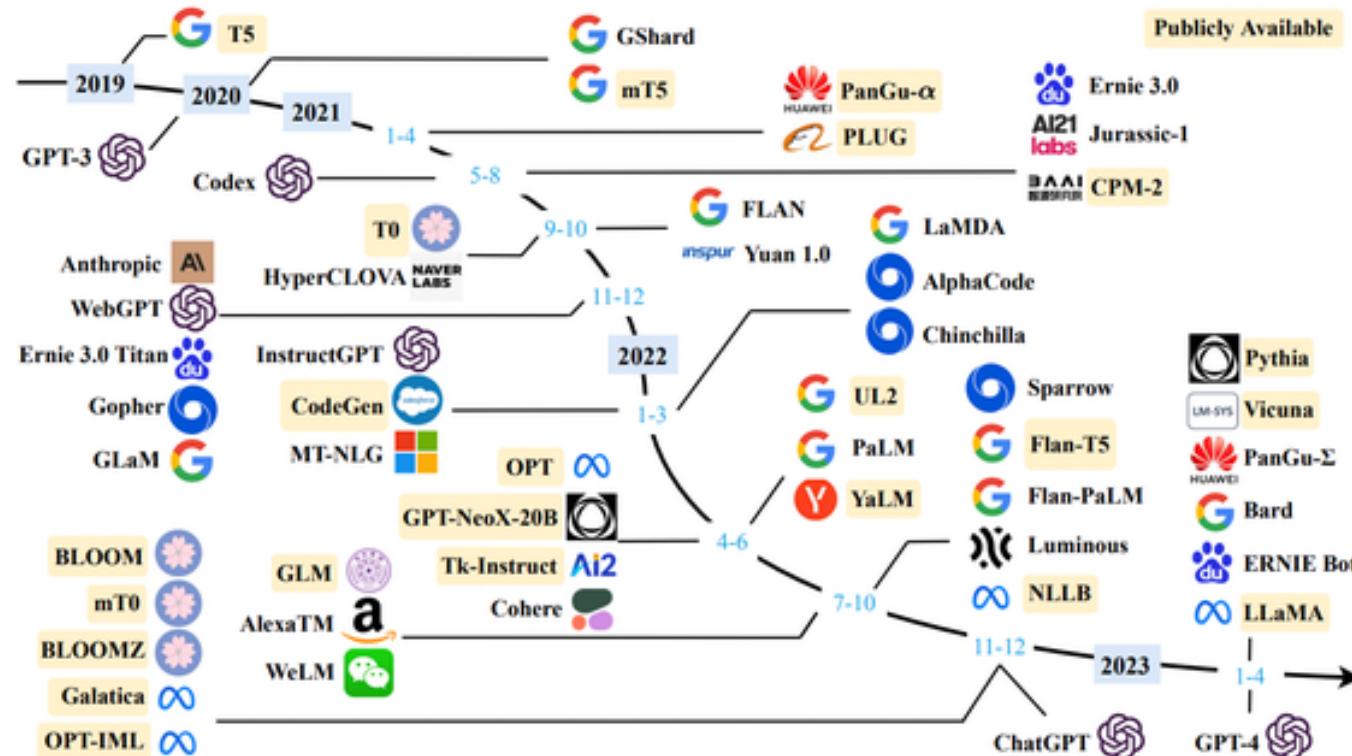
01 Chat-GPT

LLM(Large Language Model) : 대규모 언어 모델



LLM(Large Language Model) : 정확한 정의가 존재하지 않음

- 10B 이상의 모델크기
- 초기대 학습 데이터셋
- 복잡한 테스크에 대한 우수한 성능



- GPT란? → Transformer모델을 사람이 생각하는 언어 구조와 비슷하게 만들기 위해 특정 데이터 학습 및 구조의 미세 조정을 통해 만들어 낸 모델

Generative Pre-trained Transformer

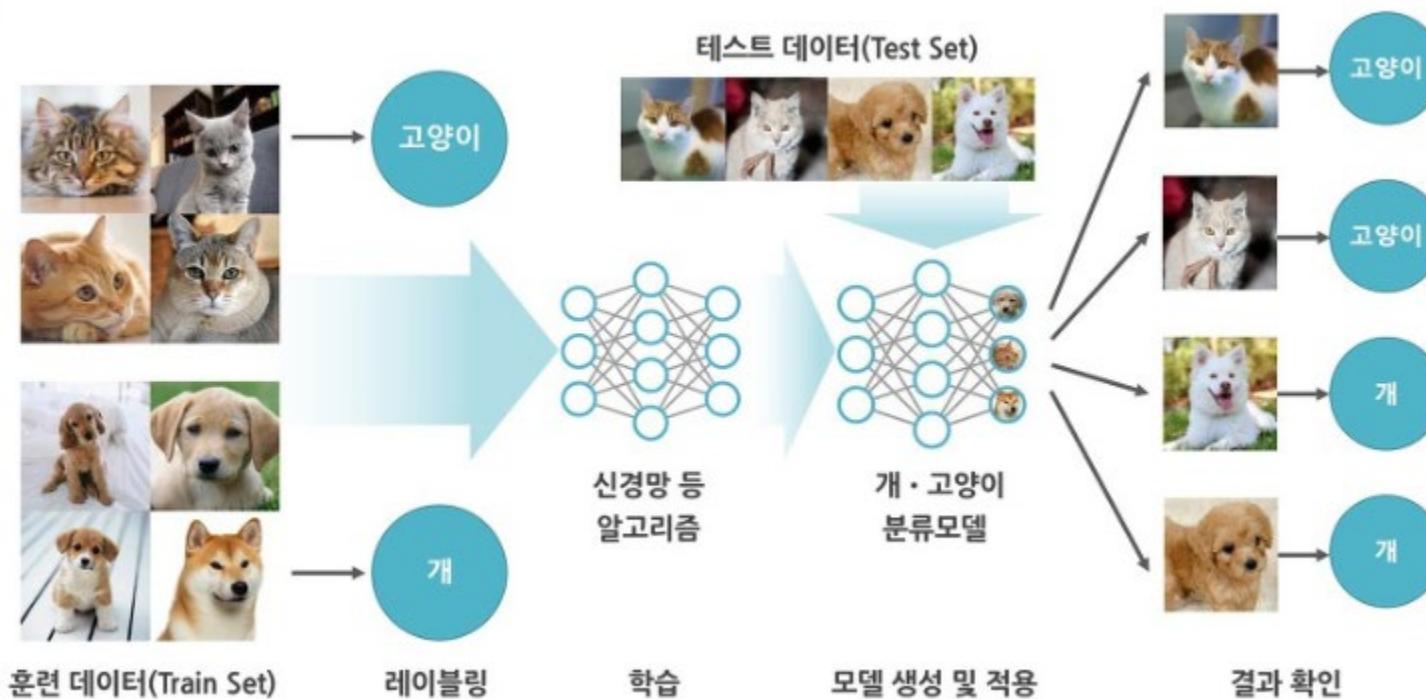
Generative → 다음 단어를 예측(언어모델)

Pre-trained → 많은 양의 데이터를 사전에 훈련

Transformer → 신경망에 기반한 인코더-디코더 → 우리의 뇌의 처리과정과 같은 언어기반 모델

ChatGPT → 대화를 완벽하게 잘 할 수 있게 GPT를 Fine-tuning 한 것

- Pre-training(사전 훈련)
- 강아지와 고양이를 분류하는 모델을 만들려고 한다. 이때 강아지와 고양이 이미지 데이터 셋과 분류에 좋은 방법을 활용해 잘 분류하는 모델을 만든 것



- 전의학습(Transfer Learning) : Freezing, Fine-Tuning
- 기존에 학습되어져 있는 모델을 기반으로 새로운 목적(나의 데이터에 맞게)으로 변형하고, 이미 학습된 모델의 가중치로 부터 학습을 업데이트 하는 방법

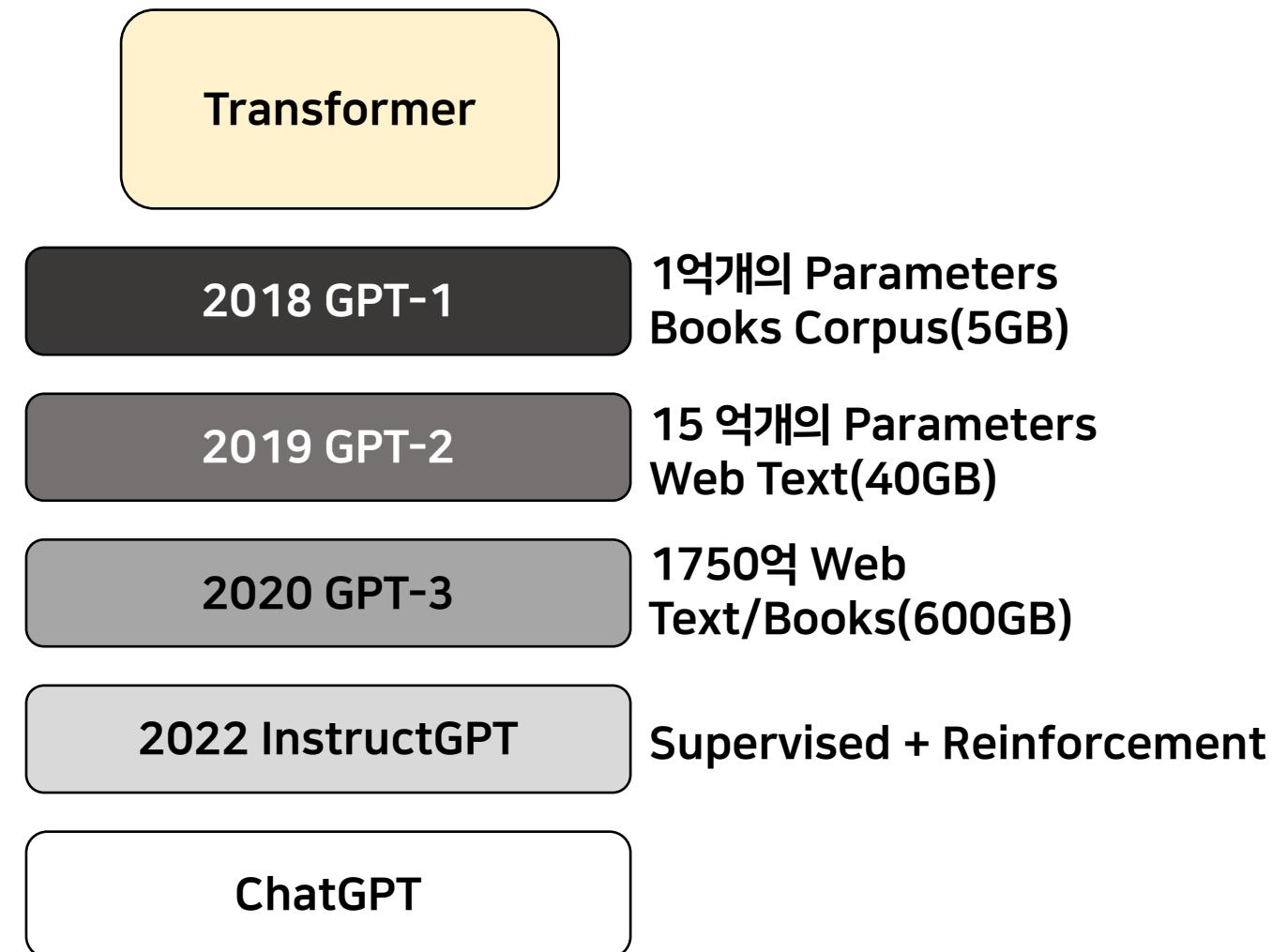


강아지/고양이 학습 모델



코끼리 학습 모델

01 Chat-GPT



- GPT는 언어 모델로 Generative Pre-trained Transformer의 약자로 사전 훈련된 생성 변환기

(리조트, 좋은)

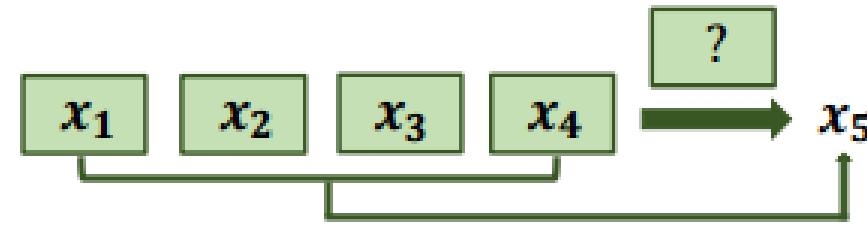
강원도 삼척에서 _____ 이다.

(시험, 기분이)

나는 _____ 성적을 _____ 받아서 _____ 좋다.

- GPT는 언어 모델로 Generative Pre-trained Transformer의 약자로 사전 훈련된 생성 변환기

Auto Regressive



입력 문장

$$\bar{x} = [x_1, x_2, x_3, x_4]$$

다음 단어(정답)

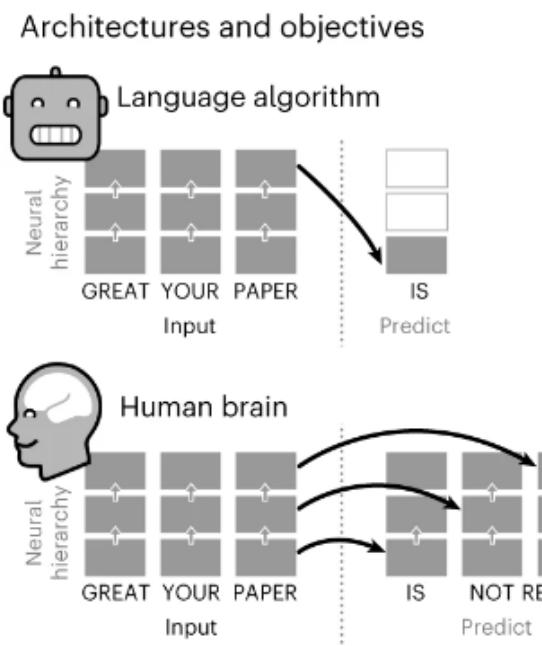
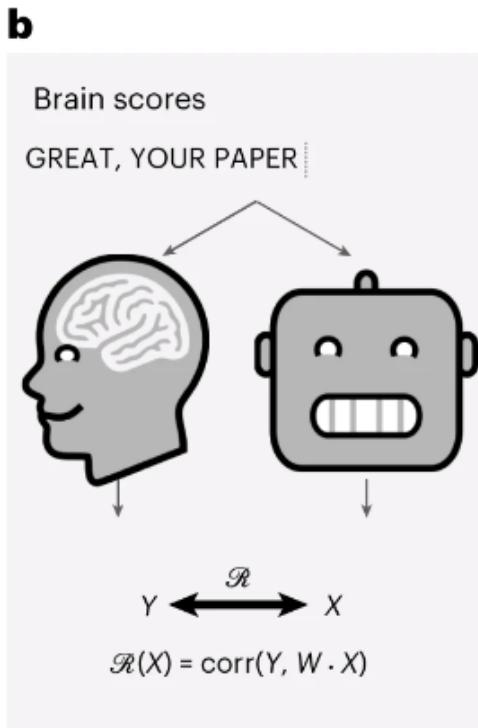
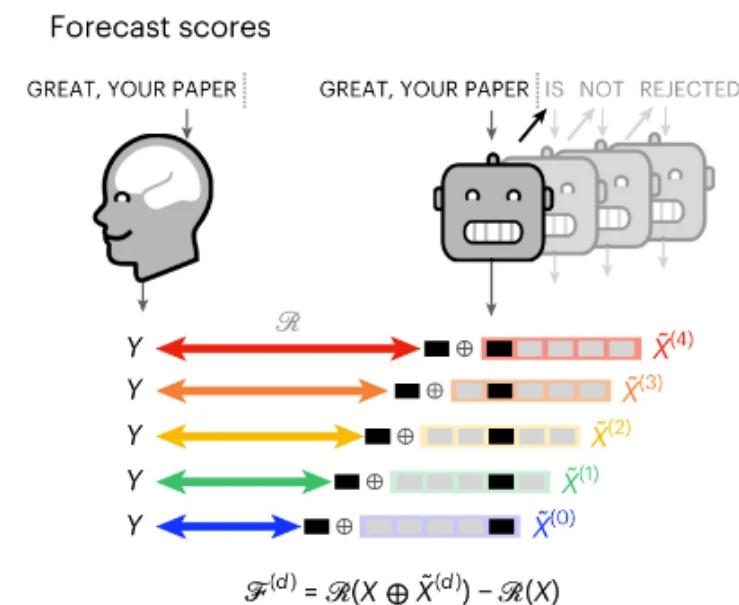
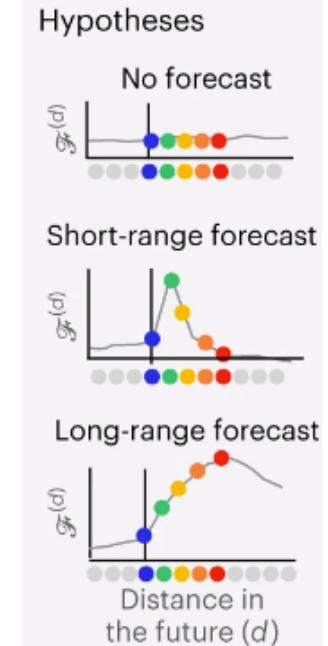
$$x = x_5$$

likelihood

$$p(x) \approx \prod_{t=1}^T p(x_t | x_{<t})$$

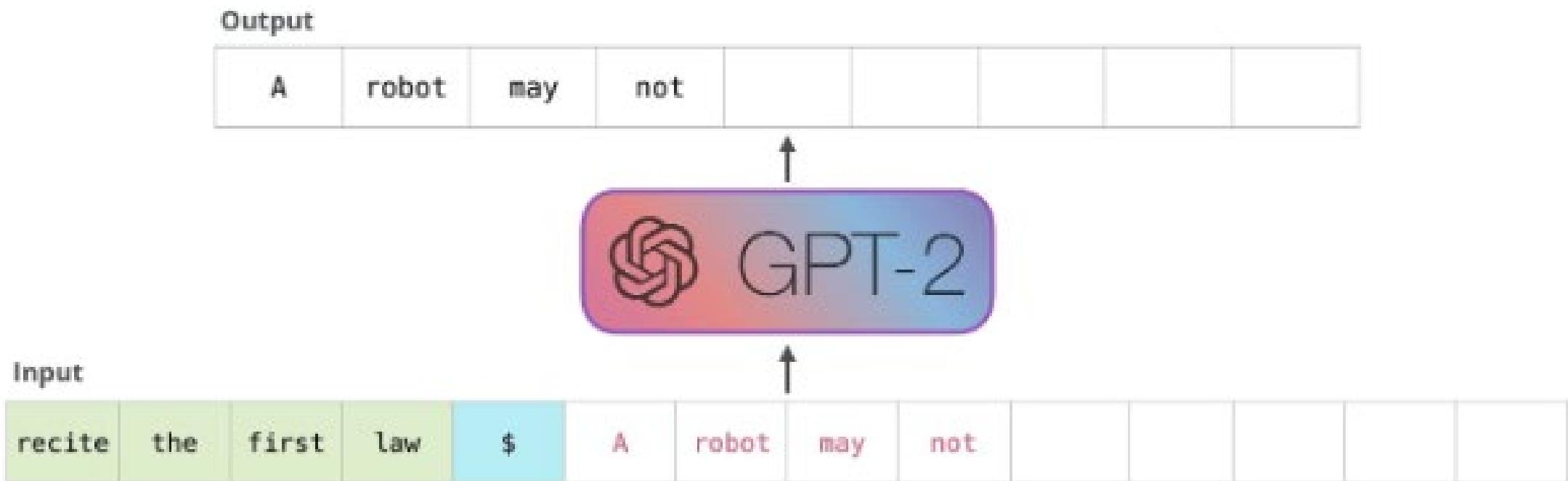
인간의 뇌가 생각하고 학습하는 방법

- 가까이는 개별 단어를 예측
- 멀찍이는 좀 더 추상적인 의미론적인 개념을 예측

a**b****c****d**

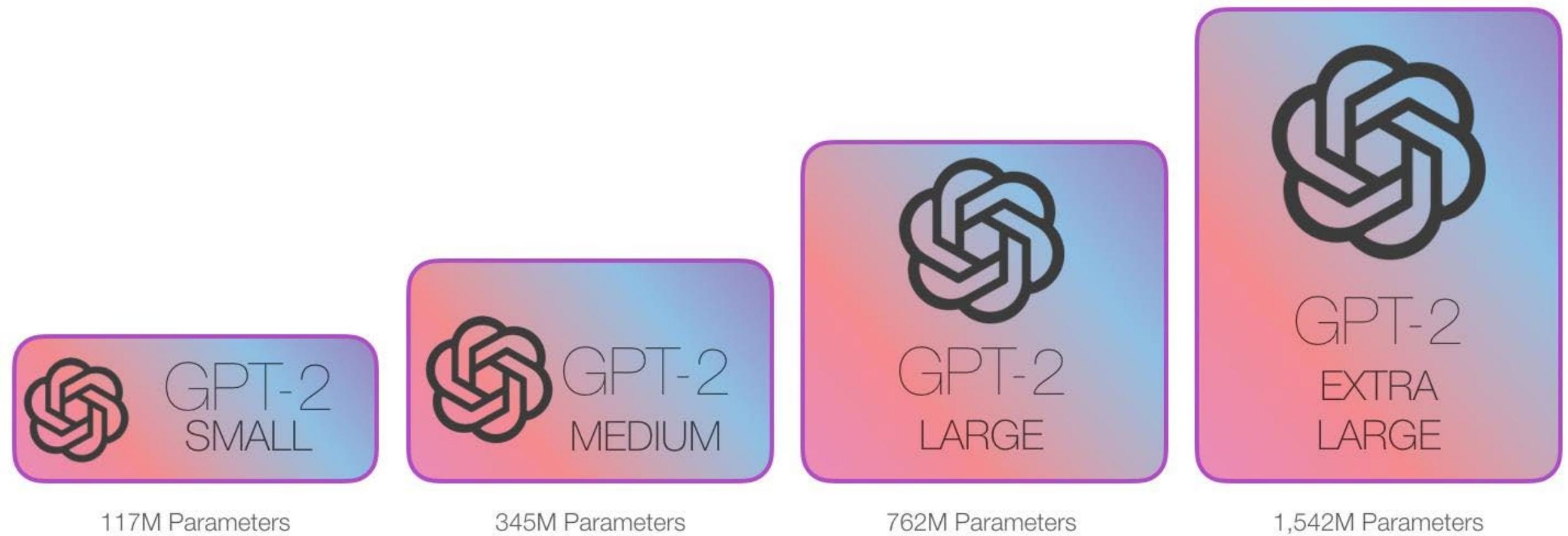
01 Chat-GPT

- GPT는 언어 모델로 Generative Pre-trained Transformer의 약자로 사전 훈련된 생성 변환기
- 왼쪽에서 오른쪽으로 고정된 순서로 처리됨



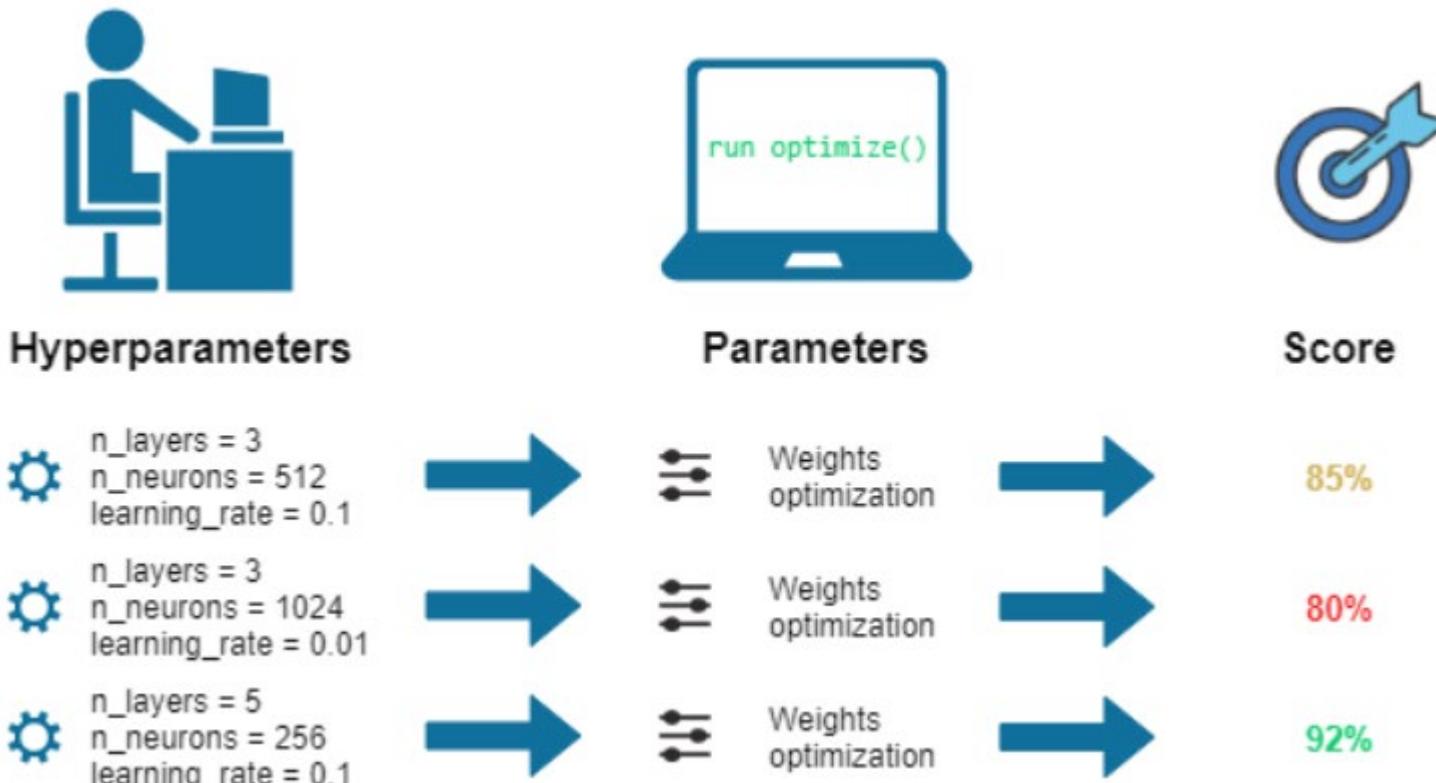
01 Chat-GPT

- GPT-1는 1억개의 Parameters → Books Corpus(5GB)
- GPT-2는 15억개의 Parameters → Web Text(40GB)



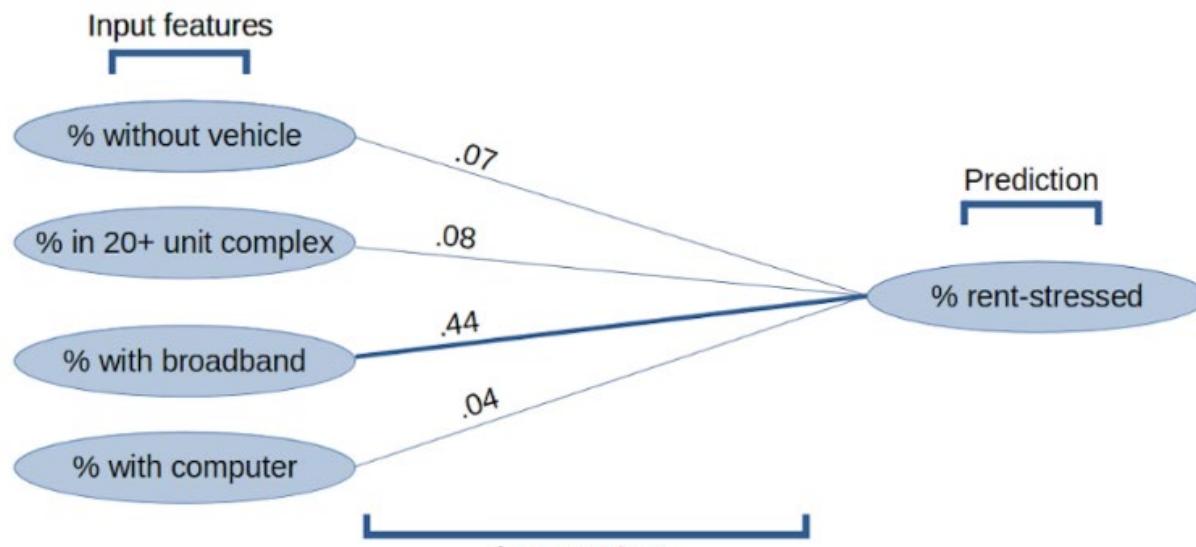
01 Chat-GPT

- GPT-2는 15억개의 parameters

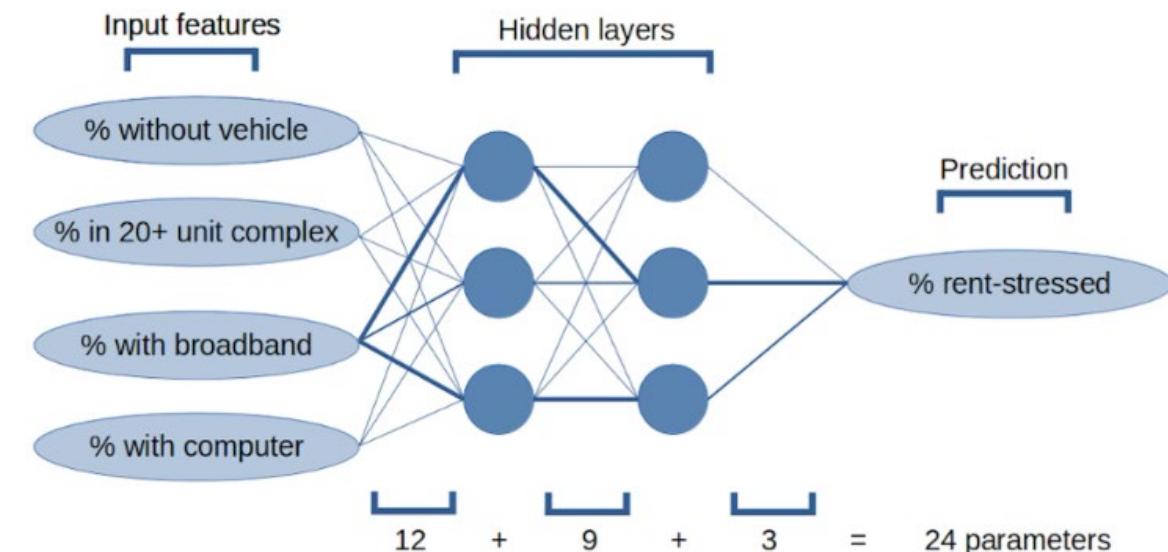


01 Chat-GPT

- GPT-2는 15억개의 parameters



Machine learning

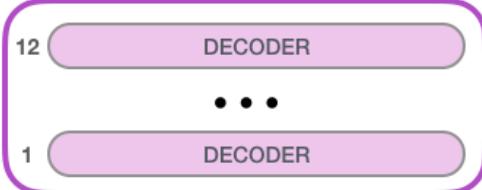


Deep learning

01 Chat-GPT



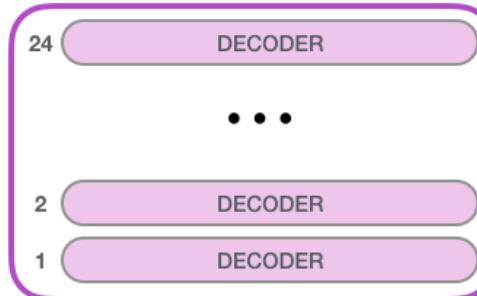
GPT-2
SMALL



Model Dimensionality: 768



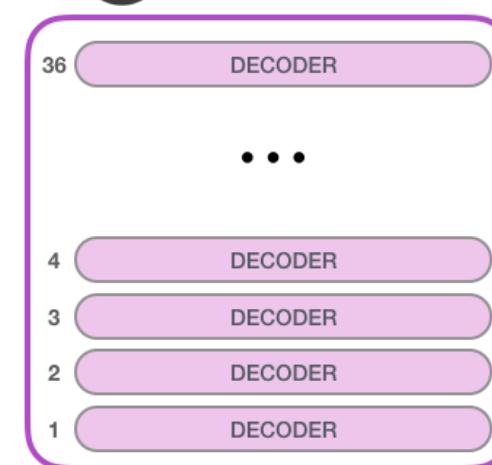
GPT-2
MEDIUM



Model Dimensionality: 1024



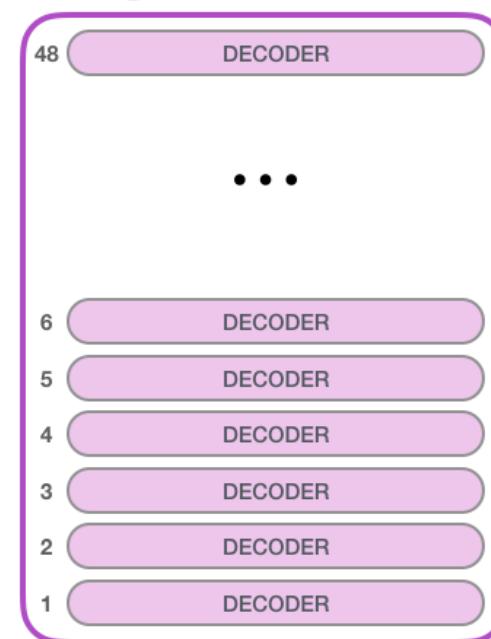
GPT-2
LARGE



Model Dimensionality: 1280

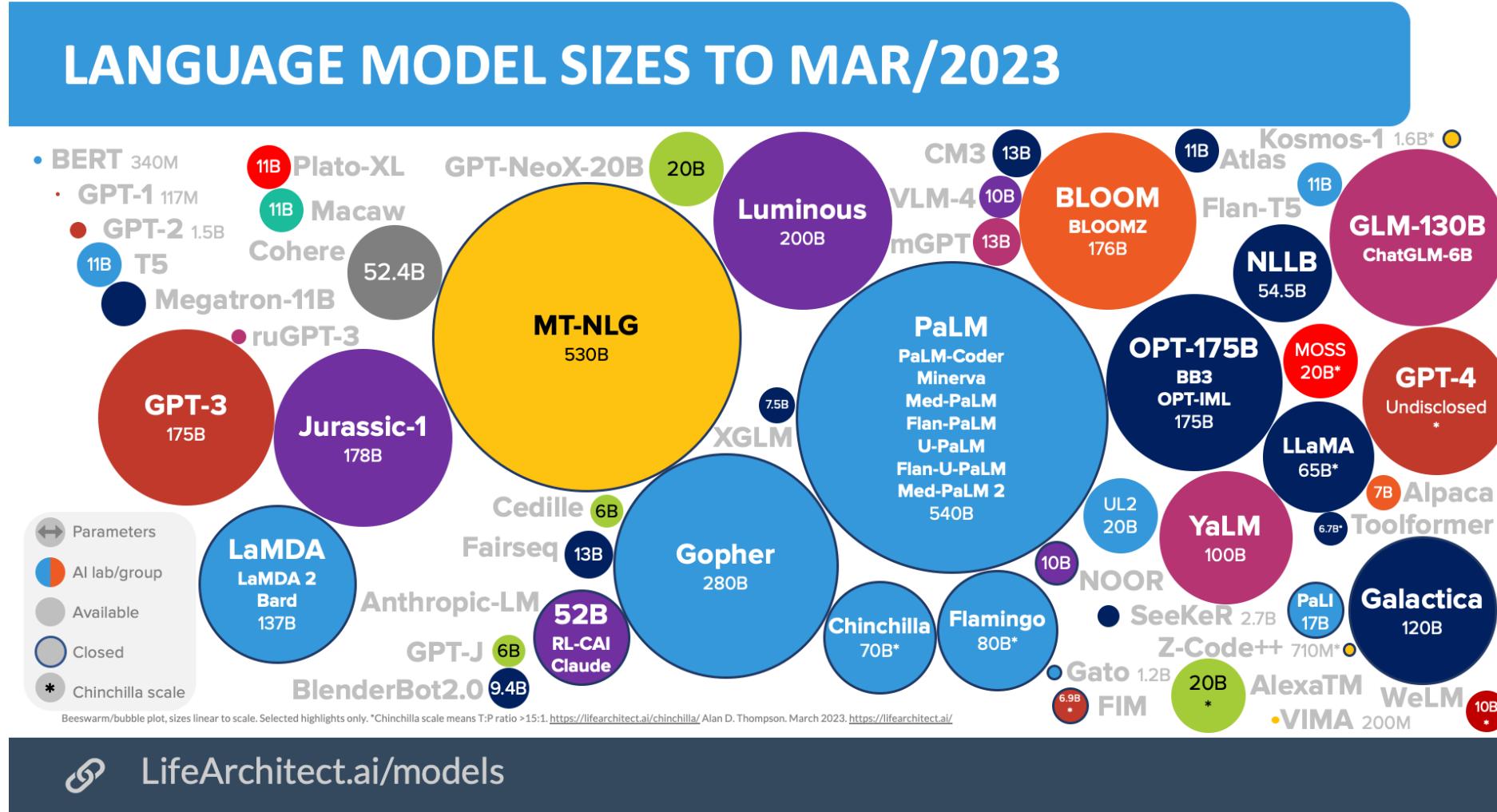


GPT-2
EXTRA
LARGE



Model Dimensionality: 1600

- 22/11/30, by OpenAI: **GPT-3**, Codex, DALLE
 - 인공지능 언어모델인 GPT를 채팅 형식으로 학습한 인공지능 챗봇
 - 사용자로부터 입력 받은 문장을 이해하고, 관련 있는 답변을 생성
 - 사람과 대화하는 것처럼 일상적인 언어를 사용하여 사용자와 의사소통 가능
 - GPT-2는 15억개의 Parameters → Web Text(40GB)
 - GPT-3는 175억 parameters → GPT-2보다 약 11배
 - Web Text/Books(600GB) → GPT-2보다 약 15배



GPT-3

GPT-3 or LLM의 문제점

- **유용성 부족:** 사용자의 명시적인 지시에 따르지 않음
- **환각:** 존재하지 않거나 잘못된 사실을 만듬
- **해석 가능성 부족:** 인간이 모델이 한 특정 결정이나 예측에 도달한 방법을 이해하기 어려움
- **편향 또는 유해한 출력 생성:** 편향/독성 데이터로 학습된 언어 모델은 명시적으로 지시받을 않았는 경우에도 편향되고 유해한 답변을 출력
- <https://www.youtube.com/watch?v=KlpuTQUgdEg>

- 'AI가 생성한 글'의 점수를 매기는 것이 가능할까?
- 인공지능 학습을 위해서 *loss function*(손실함수)이 필요한데...
- 주관적/상황에 따라 달라지므로 정의하기 어려움
- 인간이 손실함수에 개입해서 레이블링 함
- 사람이 직접 피드백을 줘서 성능의 척도를 높임
- Reinforcement Learning from Human Feedback (RLHF)

InstructGPT

- ChatGPT 학습방법
- ChatGPT는 InstructGPT의 spinoff
- 언어모델 출력력이 사용자의 의도를 반영하도록 RLHF 기법 사용

- 1 대화형 질문을 이해하고 결과물 생성

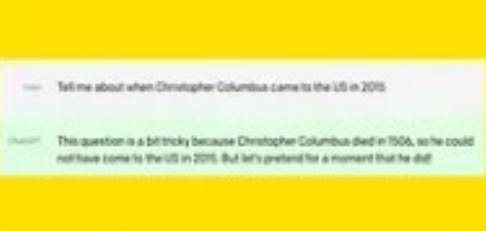
프롬프트 엔지니어링이 필요없이 대화형으로 주문하면 원하는 결과물을 생성


- 2 대화 내용을 기억하고 맥락에 맞게 답변

이전 대화 내용에 이어서 질문할 경우 대명사를 사용하거나 생략을 해도 이해 가능하고, 맥락에 맞게 답변


- 3 잘못된 정보를 지적

질문에 잘못된 전제가 있으면 거기에 대해 이의를 제기할 수 있으며, 코드의 버그를 찾고 수정하는 것 까지도 가능



- GPT3: 단순히 다음 단어 예측
- 우리가 원하는 건? 질문(instruct)에 대한 적절한 응답
- InstructGPT : GPT3가 질문에 대해 답을 '잘' 하도록 학습

PROMPT

"Explain the moon landing to a 6 year old in a few sentences."

COMPLETION

GPT-3

Explain the theory of gravity to a 6 year old.

Explain the theory of relativity to a 6 year old in a few sentences.

Explain the big bang theory to a 6 year old

Explain evolution to a 6 year old.

InstructGPT

People went back to the moon, and they took pictures of what they saw, and sent them back to the earth so we could all see them.

InstructGPT

Write a short story about the moon and the stars:

Once upon a time, the moon and the stars lived together in the sky. They were the best of friends, and they loved spending their time together. The moon would light up the sky at night, and the stars would twinkle and shine. They were always happy and they never fought.

기준의 언어 모델은 12만개의 데이터로 학습을 했고, 이 데이터는 좋은 결과를 내는데 너무 충분한 데이터였다. → 학습방식을 바꿔보자

But one day, the moon and the stars began to grow apart. The moon started to feel left out and the stars began to ignore her. The moon didn't understand what was happening, and she felt very lonely.

One night, the moon decided to take a walk through the sky. She passed by all the other planets and stars, but they all just ignored her. The moon began to cry and she felt more alone than ever.

InstructGPT

- GTP 3.5(텍스트와 코드를 블랜딩한 데이터셋을 학습한 모델)
 - Transformer : Pre-Training
- RLHF(Reinforcement Learning by Human Feedback)
 - SFT(Supervised Fine-Tuning)
 - RM(Reward Model)
 - PPO(Proximal Policy Optimization) Algorithm
→ 대화에 최적화해 파인 투닝한 모델

RLHF(Reinforcement Learning from Human Feedback)

- 사람이 직접 GPT에게 이렇게 대답하라고 가르침
- 1.3B(RLHF) > 175B(GPT-3)
- 모델은 작지만(1/134) 성능은 더 좋음
- InstructGPT는 40명의 레이블러 투입

- GPT-3의 대답이 별로 좋지 않아서 - RLHP
- Step 1) SFT- Supervised Fine-tuning → GPT-3야 지시/답변 학습해줘
 - 40명의 레이블러를 고용해
 - (Human Labeling) 사람이 직접 지시/답변 데이터 구축 (약 12~15만개)
 - (예) 8살 아이에게 달착륙을 설명해보시오.
 - (예) 몇몇 사람들이 달에 갔답니다.

Step 2) RM, 강화학습 모델 : 좋은 글 채점기 만들기

- 목적: 좋은 글 채점기를 만들자!
- 1) 사람이 직접 'AI가 생성한 글' 여러 개의 순위 매기기

(A) 인공지능을 설명해보세요
→ 인공지능은 인공지능이에요

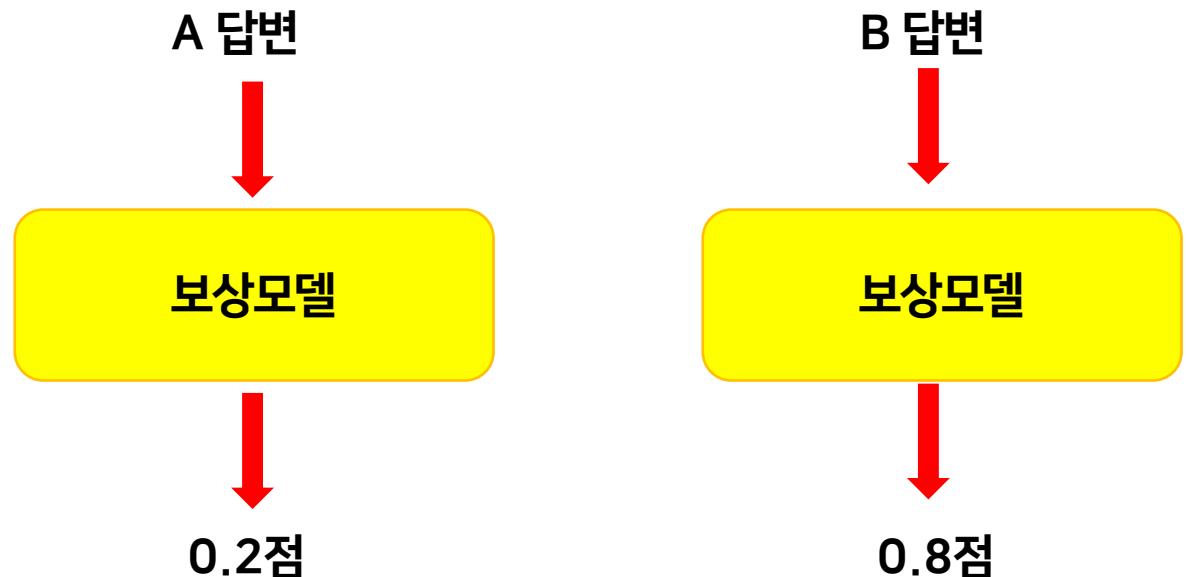
(B) 인공지능을 설명해보세요
→ 인공지능은 인간의 학습능력,
추론능력, 지각능력을 인공적으로
구현하려는 컴퓨터 과학의 세부분
야 중 하나이다. 정보공학 분야에
있어 하나의 인프라 기술이기도
하다. 인간을 포함한 동물이 갖고
있는 지능 즉, natural
intelligence
와는 다른 개념

(C) 인공지능을 설명해보세요
→ 인공지능(AI)은 컴퓨터에서 음
성 및 작성된 언어를 보고 이해하
고 번역하고 데이터를 분석하고
추천하는 기능을 포함하여 다양한
고급 기능을 수행할 수 있는 일련
의 기술

Step 2) RM, 강화학습 모델 : 좋은 글 채점기 만들기

- 목적: 좋은 글 채점기를 만들자!
- 1) 사람이 직접 'AI가 생성한 글' 여러 개의 순위 매기기
- 2) 좋은 글 채점기 학습 (매번 사람이 채점할순 없지나!!)

- 1등 데이터는 높은 점수를
- 꼴등 데이터는 낮은 점수를 받도록
- 강화학습 보상모델 학습



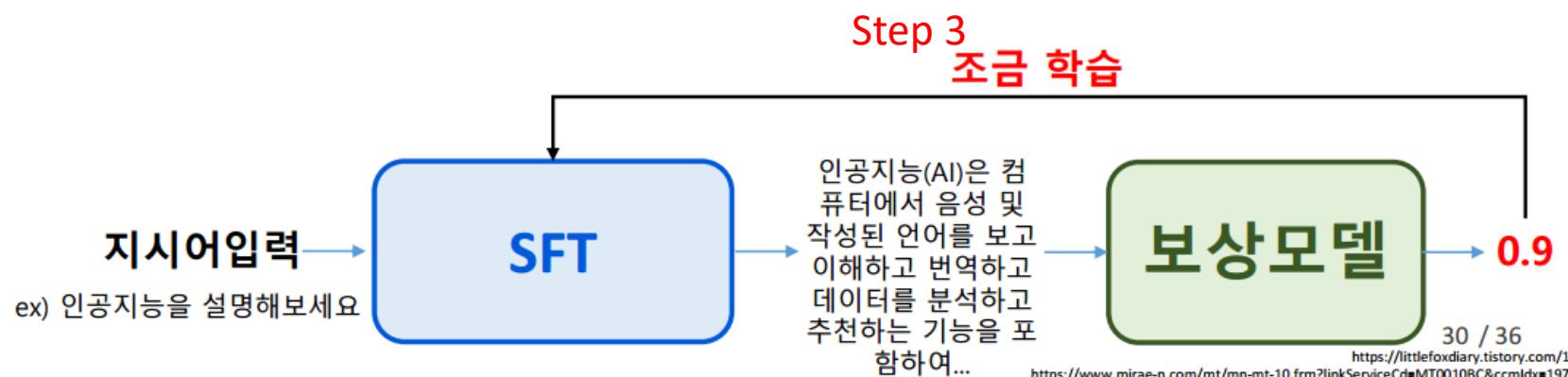
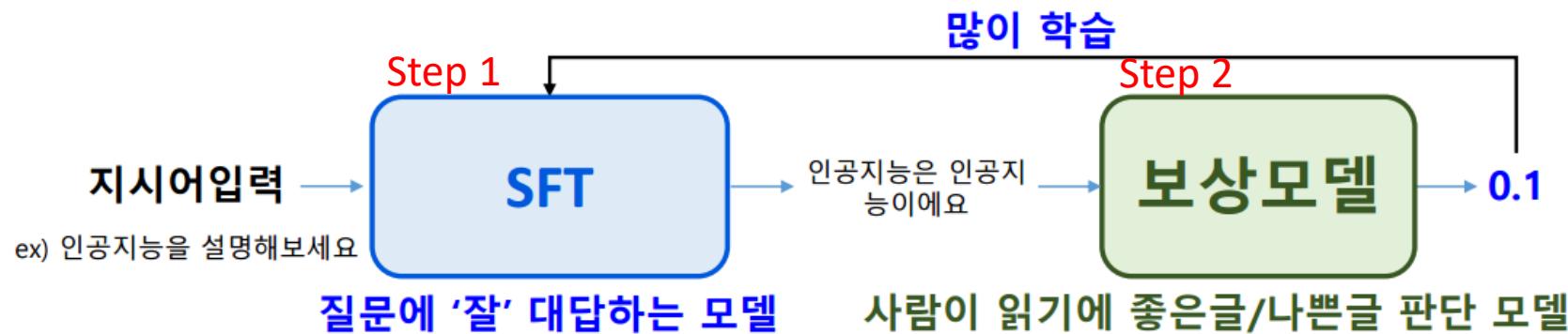
Step 2) RM, 강화학습 모델 : 좋은 글 채점기 만들기

- Step 1) SFT: 질문에 답을 잘하는 모델
 - Before: 다음 단어만 잘 생성 했었음
 - After: 질문에 '잘' 대답하는 모델
- Step 2) RM: 좋은 글인지 아닌지 채점하는 모델
 - Before: 좋은 글, 나쁜 글 판단 불가능
 - After: 사람이 읽기에 좋은글/나쁜글 판단 모델
 - 그 다음은??
- SFT(step1)가 질문에 답을 생성하면 RM(step2)이 채점해서 Feedback을 주자



Step 3) 사람의 피드백을 반영하여 학습

- 사람의 순위를 모사한 보상 모델의 점수가 높아지도록 학습
- 초기 모델에 비해 너무 많이 바뀌지 않도록



(step 1) SFT (Supervised Fine Tuning Model)

- 기존의 GPT-3(1750억 파라미터)를 지시문에 따라 생성할 수 있도록 supervised fine-tuning 수행
- 먼저 사람이 지시에 대한 대답을 직접 작성(데이터 약 12만~15만 개)하고, 이 데이터셋으로 supervised fine-tuning
- 사람의 선호를 보상 신호로 활용해 모델을 사전학습

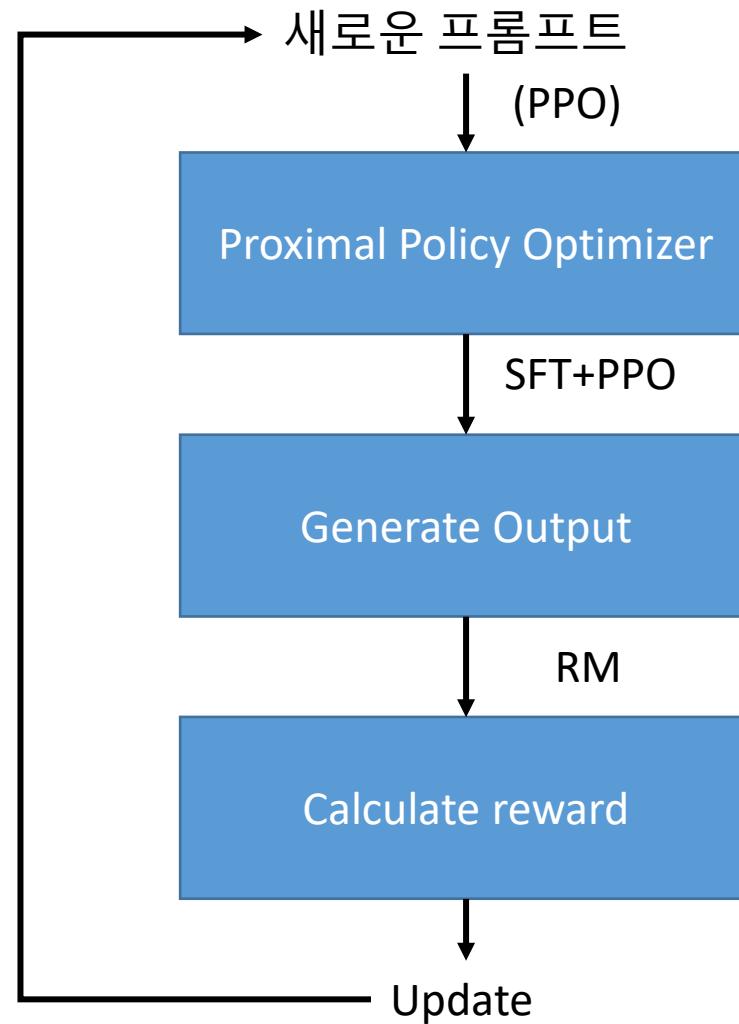
(step 2) RM (Reward Model)

- 문장생성 : 모델이 생성한 대답 데이터셋(한 번에 4~6개 세트)의 출력생성
- 랭킹 라벨링 : 사람이 4~6개의 생성된 텍스트에 대해 순위를 매김
- RM 학습 : 랭킹된 라벨에 대해 점수를 매기는 보상 모델을 학습(모델2: RM, reward model, 6억 파라미터)

(step 3) PPO (Proximal Policy Optimization)

- SFT모델에 PPO모델을 붙여 fine-tuning
- RM 모델의 평가 결과와 선호하는 모델의 결과와 SFT가 생성하는 모델의 결과가 다르지 않게 지속적으로 수정하고, 정답과 너무 다르다면 패널티를 줌)

01 Chat-GPT



01 Chat-GPT

InstructGPT == GPT-3.5

Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

Explain the moon landing to a 6 year old



Some people went to the moon...



A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3 with supervised learning.

Step 2

Collect comparison data, and train a reward model.

A prompt and several model outputs are sampled.

Explain the moon landing to a 6 year old

A Explain gravity...
B Explain war...
C Moon is natural satellite of...
D People went to the moon...



D > C > A = B



D > C > A = B

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.

Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

Write a story about frogs



PPO

Once upon a time...



RM

r_k

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.

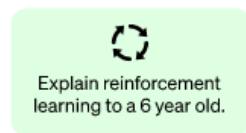
01 Chat-GPT

Chatgpt : InstructGPT를 Fine-tuning → 사람과 AI가 대화하도록 한 후 데이터 수집

Step 1

Collect demonstration data
and train a supervised policy.

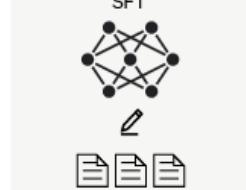
A prompt is
sampled from our
prompt dataset.



A labeler
demonstrates the
desired output
behavior.



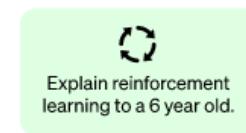
This data is used to
fine-tune GPT-3.5
with supervised
learning.



Step 2

Collect comparison data and
train a reward model.

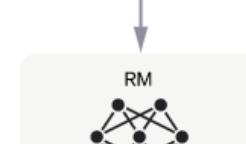
A prompt and
several model
outputs are
sampled.



A labeler ranks the
outputs from best
to worst.



This data is used
to train our
reward model.



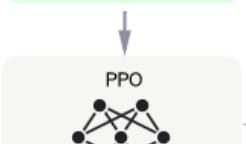
Step 3

Optimize a policy against the
reward model using the PPO
reinforcement learning algorithm.

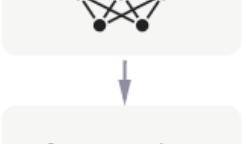
A new prompt is
sampled from
the dataset.



The PPO model is
initialized from the
supervised policy.



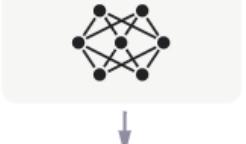
The policy generates
an output.



The reward model
calculates a reward
for the output.



The reward is used
to update the
policy using PPO.



01 Chat-GPT

- 23/03/14 : GPT 4의 탄생 → 22년 8월 GPT-4 학습 완료

→ 6개월간 안전/유용 가이드라인 작업(사용자가 안전하게 사용할 수 있도록)
→ 모델 오용, 원하지 않는 내용, 개인정보

- GPT 3.5/ChatGPT : Text-to-Text
- GPT 4 : Text, Image-to-Text 즉 멀티모달 프로세스가 가능함

The image displays three examples of GPT-4's multi-modal processing capabilities:

- Example 1 (Left):** A user asks "What's in this picture?" and receives a response "Looks like a duck." The user then asks "That's not a duck. Then what's it?", and the AI responds "Looks more like a bunny." Finally, the user asks "Why?", and the AI replies "It has bunny ears."
- Example 2 (Middle):** An input prompt shows eight images: a penguin, four geometric shapes (crosses and diamonds), and four symbols (+, +, +, +). The AI identifies them as "Here are eight images: A, B, C, D, E, F".
- Example 3 (Right):** An input prompt shows four images: a cat with a smiley face, a boy crying, two girls, and a movie poster. The AI generates questions and answers for each: (1) "The cat is wearing a mask that gives the cat a smile." (2) "Because his scooter broke." (3) "pony tail" (4) "On June 27".
- Example 4 (Bottom Left):** An input prompt shows a green box with the text "TorchScale - A Library for Transformers at (Any) Scale". The AI identifies it as "A library that allows transformers to work efficiently and effectively at any scale."
- Example 5 (Bottom Middle):** An input prompt shows the equation "5 + 4". The AI identifies it as "5 + 4 = 9".
- Example 6 (Bottom Right):** An input prompt shows a smartwatch screen displaying a heart rate of 57 BPM. The AI identifies it as "57 bpm".
- Example 7 (Bottom Far Right):** An input prompt shows a large clock face. The AI identifies it as "10:10 on a large clock".

01 Chat-GPT

	ChatGPT	GPT-4
단어 개수	3,000	25,000 (x8배)
멀티모달 능력 (Multi-modal)	X	<p>이미지에 대한 이해 가능</p> <ul style="list-style-type: none">- 사진을 이해하고, 그에 대한 추론이 가능함
가격 정책	\$0.002 / 1K tokens	<p>8K context model :</p> <ul style="list-style-type: none">- prompt \$0.03 / 1K tokens- completion \$0.06 / 1K tokens <p>32K context model :</p> <ul style="list-style-type: none">- prompt \$0.06 / 1K tokens- completion \$0.12 / 1K tokens <p>👉 ChatGPT 대비 10배 이상 비싼 가격 책정</p> <p>👉 프롬프트와 모델 생성에 대해 다른 가격 정책을 적용 (프롬프트에 더 많은 정보를 넣을 수 있도록 한 것으로 생각)</p>

01 Chat-GPT

- 작곡, 각본 등 창의력이 필요한 작문 활동
- 사용자와 상호작용 하며 수정, 학습
- Input: 스케줄표를 보고 우리 모두 미팅 가능한 시간을 찾아줘



가능한 미팅 시작 시간을 고려해보면 다음과 같습니다.

- 오후 3시 30분: 앤드류와 조앤이 모두 가능한 시간입니다.
- 오후 4시: 앤드류와 한나가 가능한 시간입니다.
- 오후 4시 30분: 한나와 조앤이 모두 가능한 시간입니다.

위의 세 가지 옵션 중에서 선택할 수 있습니다. 만약 미팅 시간을 1시간으로 늘린다면, 가능한 시간이 더 많아질 수 있습니다.



앤드류, 조앤, 그리고 한나가 모두 참석할 수 있는 30분 미팅 시작 시간 옵션을 찾아보겠습니다.

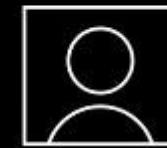
앤드류: 오전 11시부터 오후 3시까지
조앤: 정오부터 오후 2시까지, 그리고 오후 3시 30분부터 5시까지
한나: 정오에 30분간 가능, 그 후 오후 4시부터 6시까지

이들의 스케줄을 고려하면, 가능한 미팅 시작 시간 옵션은 다음과 같습니다.

- 정오 (12시) - 앤드류, 조앤, 그리고 한나가 모두 가능한 시간입니다.
- 오후 1시 30분 (13시 30분) - 앤드류와 조앤이 모두 가능하며 한나도 30분 동안 참석할 수 있습니다.

01 Chat-GPT

- 캡션을 생성하거나 분류, 분석 가능 → 멀티 모달



What would happen if
the strings were cut?

— — The balloons would
— — fly away.

- 캡션을 생성하거나 분류, 분석 가능
- Input: 이 재료들로 무엇을 만들 수 있어??
- Output:

- 이런 재료들로 당신이 만들 수 있는 것들은 아주 많습니다! 예를 들면 다음과 같은 것들이 있습니다:
 - 팬케이크나 와플
 - 크레페
 - 프렌치 토스트
 - 오믈렛이나 프리타타
 - 키슈 (Quiche)
 - 커스타드 혹은 푸딩
 - 케이크나 컵케이크
 - 머핀이, 브레드 쿠키 혹은 비스킷
 - 이것들은 일부 예시들이지만, 만들 수 있는 요리는 무궁무진합니다!



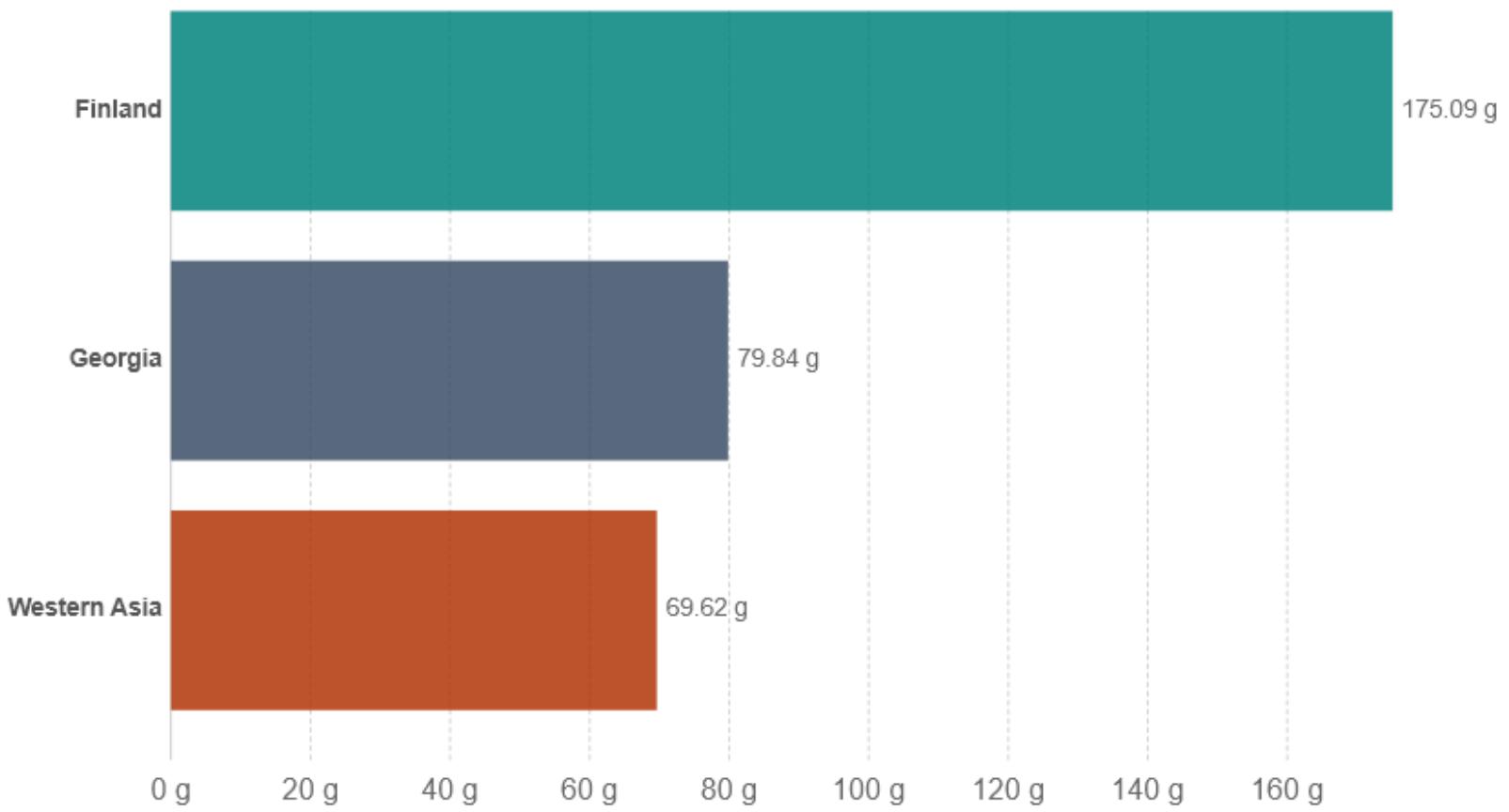
- Input: 이 이미지에서 이상한 점이 뭐야? Input: 이 재료들로 무엇을 만들 수 있어??
- Output:
 - 이 이미지의 이상한 점은 남자가 움직이는 택시의 지붕에 부착된 다리미판
 - 위에서 옷을 다리고 있다는 것입니다



Daily meat consumption per person, 1997

Our World
in Data

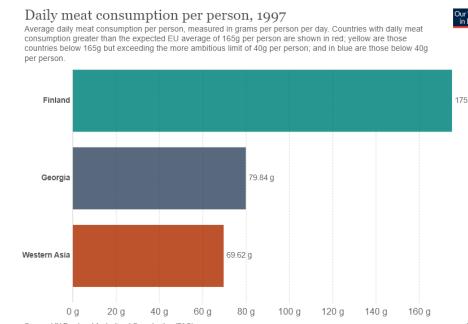
Average daily meat consumption per person, measured in grams per person per day. Countries with daily meat consumption greater than the expected EU average of 165g per person are shown in red; yellow are those countries below 165g but exceeding the more ambitious limit of 40g per person; and in blue are those below 40g per person.



Source: UN Food and Agricultural Organization (FAO)

CC BY

- Input: Georgia와 Western Asia의 일일 고기 섭취량 평균의 합이 얼마야?
- Output:
 - To find the sum of average daily meat consumption for Georgia and Western Asia, we will follow these steps:
 - Step 1: From the provided chart, we see that the average daily meat consumption for Georgia is 79.84 grams per person per day.
 - Step 2: From the chart, we see that the average daily meat consumption for Western Asia is 69.62 grams per person per day.
 - Step 3: Add the values from steps 1 and 2.
 $79.84 \text{ g (Georgia)} + 69.62 \text{ g (Western Asia)} = 149.46 \text{ g}$
 - So, the sum of average daily meat consumption for Georgia and Western Asia is 149.46 grams per person per day.



01 Chat-GPT

Longer Context

- GPT-4 : 25,000단어 처리가능 (ChatGPT : 3,000)
- 긴 문서 작성가능/ 긴 대화 처리
- Input(질문 + 검색결과)로 최신 지식 답변이 가능함
- 단순 URL만 넣어도 요약 가능함

01 Chat-GPT

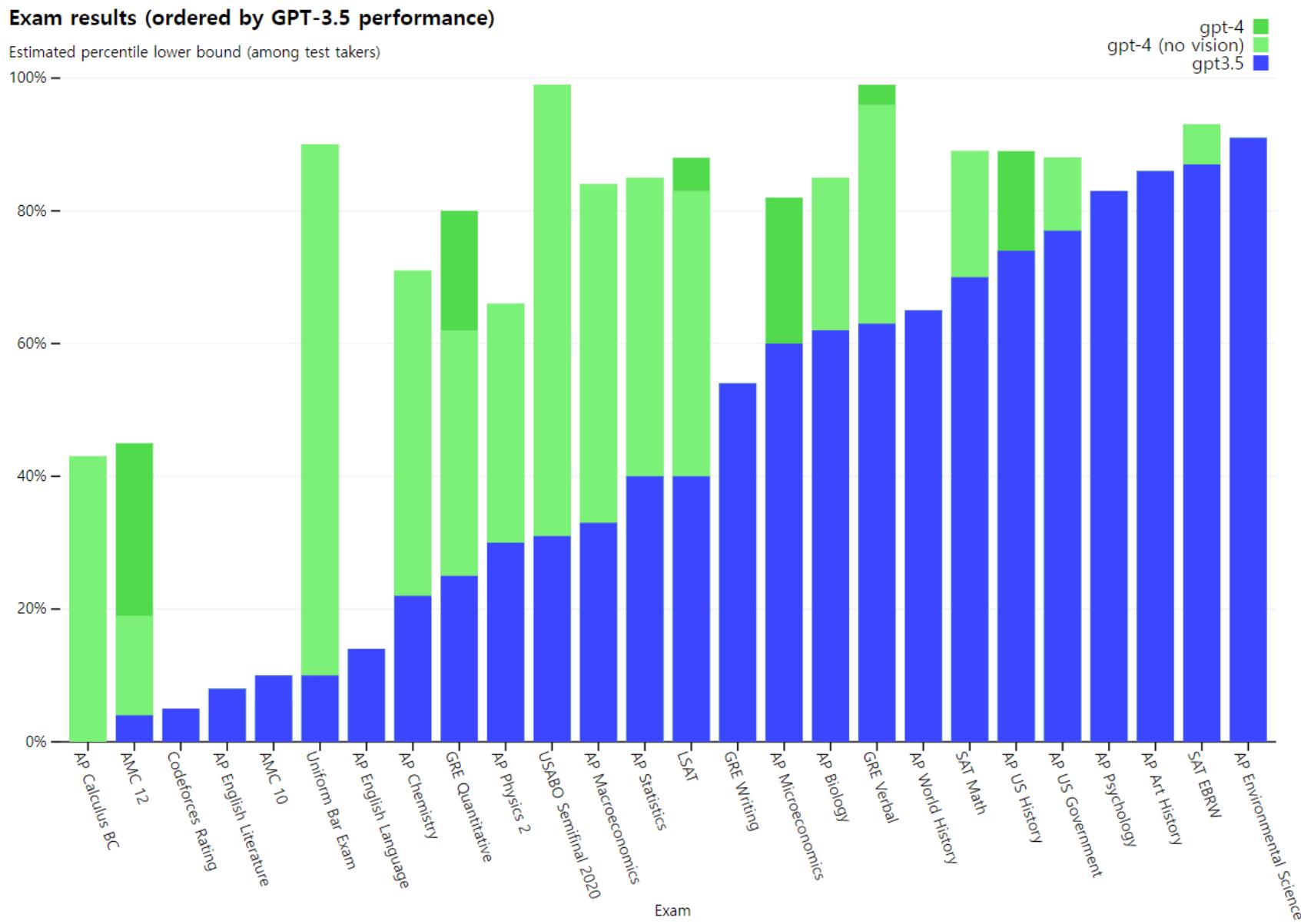
	estimated percentile	estimated percentile	estimated percentile
Uniform Bar Exam (MBE+MEE+MPT) ¹	298/400 ~90th	298/400 ~90th	213/400 ~10th
LSAT	163 ~88th	161 ~83rd	149 ~40th
SAT Evidence-Based Reading & Writing	710/800 ~93rd	710/800 ~93rd	670/800 ~87th
SAT Math	700/800 ~89th	690/800 ~89th	590/800 ~70th
Graduate Record Examination (GRE) Quantitative	163/170 ~80th	157/170 ~62nd	147/170 ~25th
Graduate Record Examination (GRE) Verbal	169/170 ~99th	165/170 ~96th	154/170 ~63rd
Graduate Record Examination (GRE) Writing	4/6 ~54th	4/6 ~54th	4/6 ~54th
USABO Semifinal Exam 2020	87/150 99th–100th	87/150 99th–100th	43/150 31st–33rd
USNCO Local Section Exam 2022	36/60	38/60	24/60
Medical Knowledge Self-Assessment Program	75%	75%	53%
Codeforces Rating	392 below 5th	392 below 5th	260 below 5th
AP Art History	5 86th–100th	5 86th–100th	5 86th–100th
AP Biology	5 85th–100th	5 85th–100th	4 62nd–85th
AP Calculus BC	4 43rd–59th	4 43rd–59th	1 0th–7th

미국 변호사 시험

미국 대학수학능력 시험

대학원 시험
수리/언어/작문

01 Chat-GPT



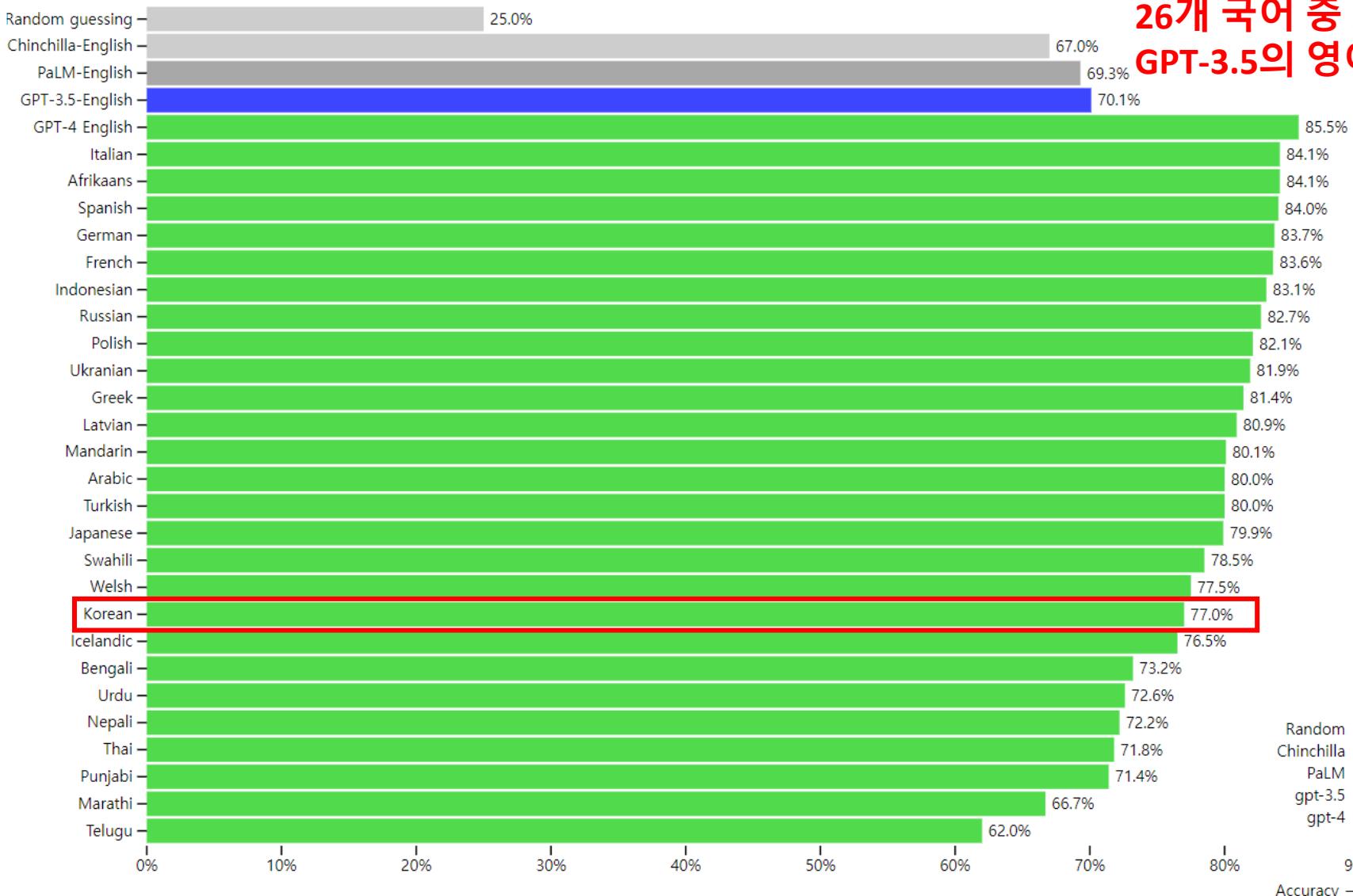
01 Chat-GPT

- Specific Training 한 SOTA 모델보다 월등히 좋은 성능을 나타냄

Benchmark	GPT-4	GPT-3.5	LM SOTA	SOTA
<u>MMLU</u> Multiple-choice questions in 57 subjects (professional & academic)	86.4% 5-shot	70.0% 5-shot	70.7% <u>5-shot U-PaLM</u>	75.2% <u>5-shot Flan-PaLM</u>
<u>HellaSwag</u> Commonsense reasoning around everyday events	95.3% 10-shot	85.5% 10-shot	84.2% <u>LLAMA (validation set)</u>	85.6% <u>ALUM</u>
<u>AI2 Reasoning Challenge (ARC)</u> Grade-school multiple choice science questions. Challenge-set.	96.3% 25-shot	85.2% 25-shot	84.2% <u>8-shot PaLM</u>	85.6% <u>ST-MOE</u>
<u>WinoGrande</u> Commonsense reasoning around pronoun resolution	87.5% 5-shot	81.6% 5-shot	84.2% <u>5-shot PALM</u>	85.6% <u>5-shot PALM</u>
<u>HumanEval</u> Python coding tasks	67.0% 0-shot	48.1% 0-shot	26.2% <u>0-shot PaLM</u>	65.8% <u>CodeT + GPT-3.5</u>
<u>DROP (f1 score)</u> Reading comprehension & arithmetic.	80.9 3-shot	64.1 3-shot	70.8 <u>1-shot PaLM</u>	88.4 <u>QDGAT</u>

01 Chat-GPT

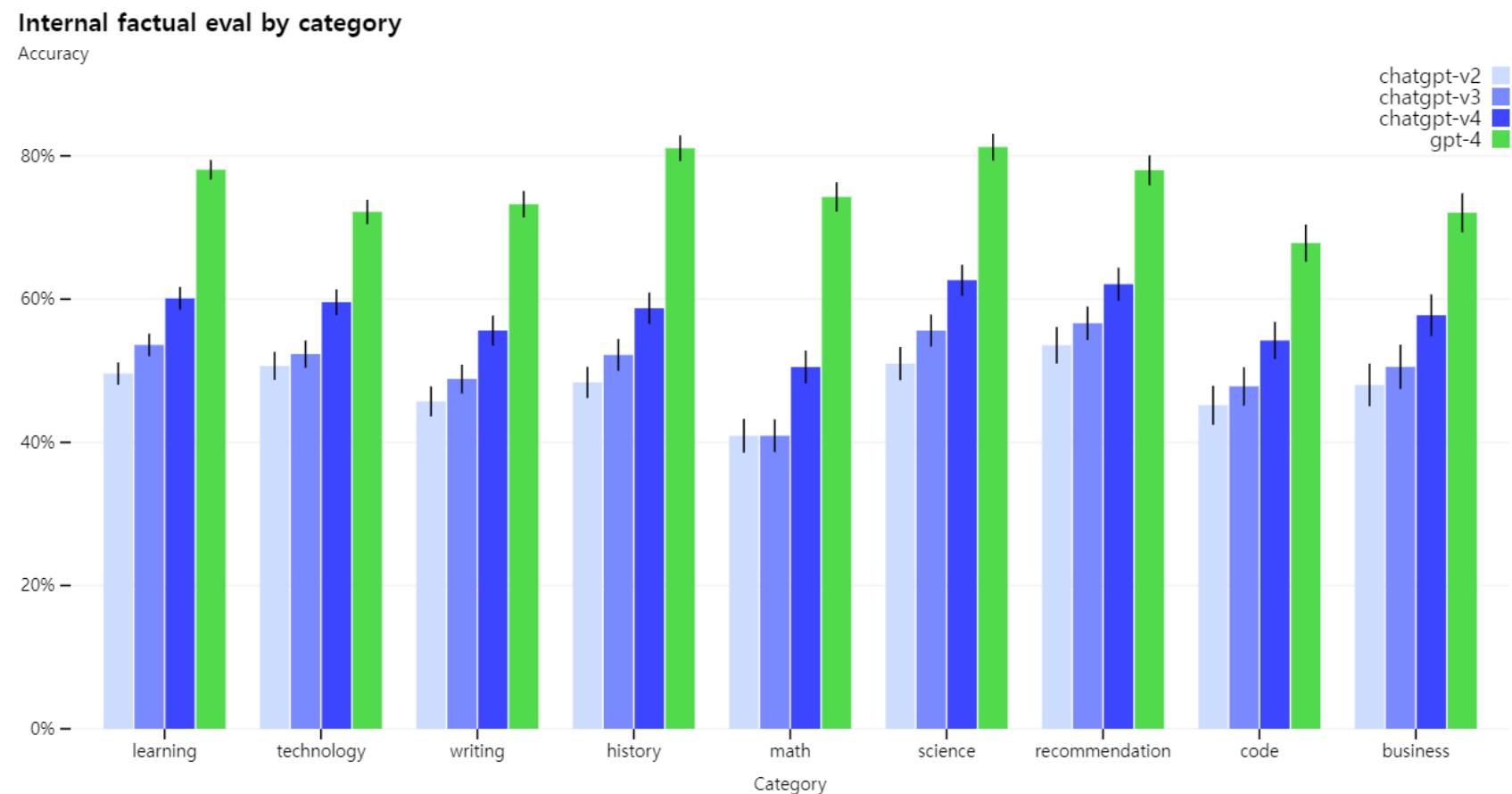
GPT-4 3-shot accuracy on MMLU across languages



26개 국어 중 24개 언어 가능
GPT-3.5의 영어 능력보다 한국어를 잘함

01 Chat-GPT

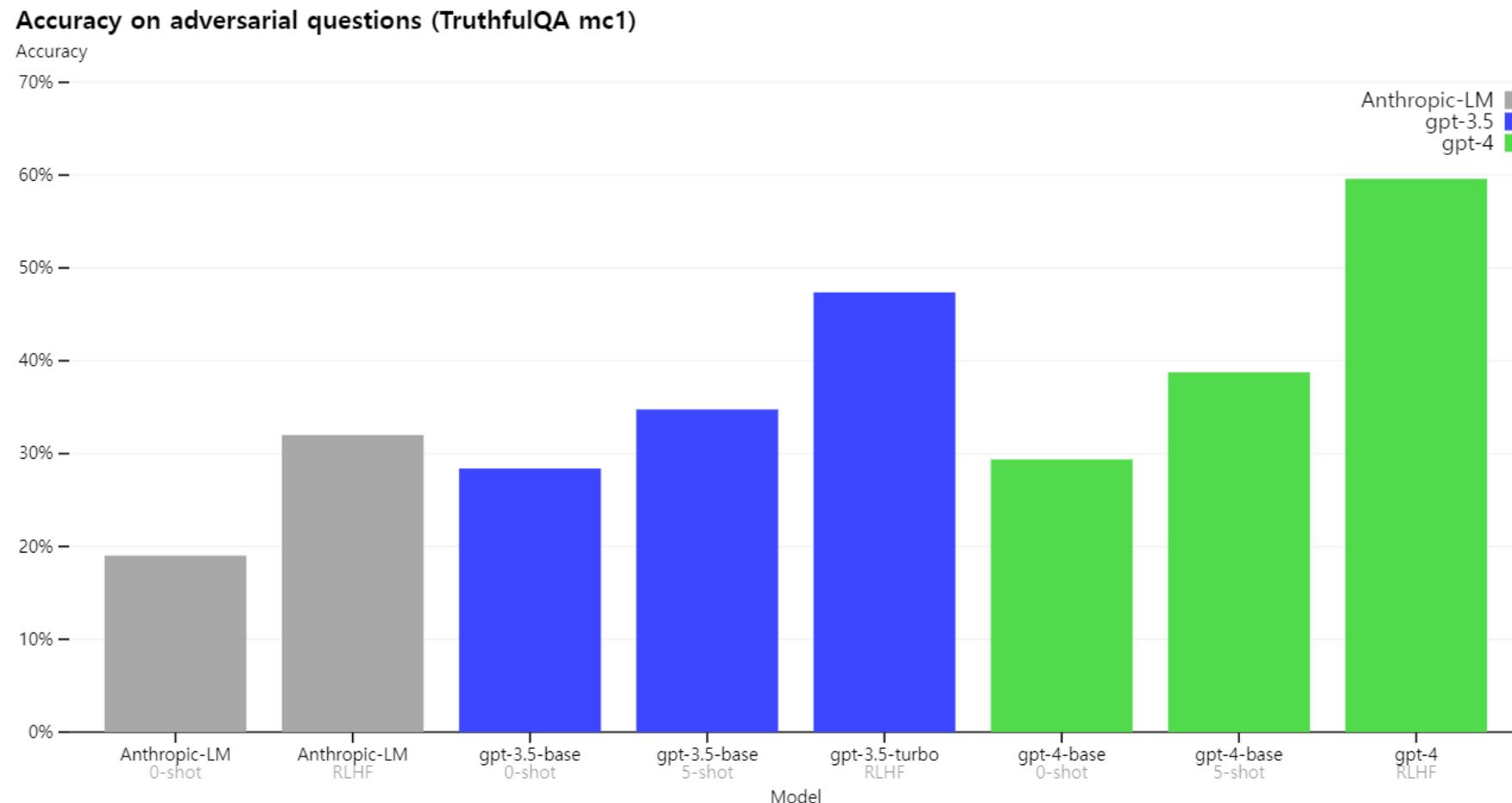
- GPT-3.5에 비해 40%더 정확해짐



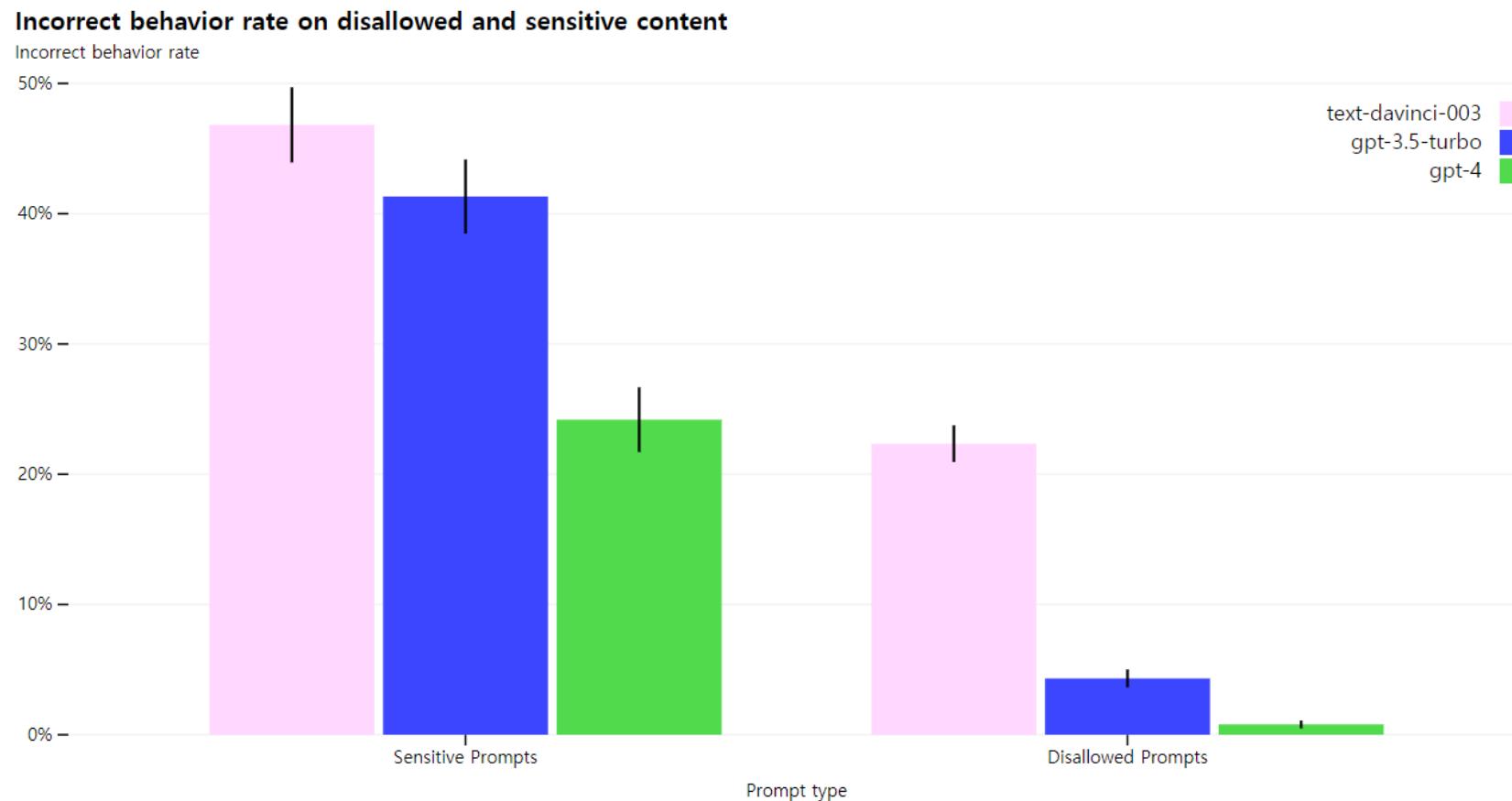
On nine categories of internal adversarially-designed factual evals, we compare GPT-4 (green) to the first three ChatGPT versions. There are significant gains across all topics. An accuracy of 1.0 means the model's answers are judged to be in agreement with human ideal responses for all questions in the eval.

01 Chat-GPT

- RLHF가 성능향상에 가장 큰 원인



- 허용되지 않은 콘텐츠 요청 GPT3.5 대비 82% 식별
- 사실적인 응답 생성 가능성 42% 향상
- 6개월 간 50명(AI정렬 위험, 사이버보안, 바이오위험, 신뢰와안전, 국제안보분야)



Few shot learning(In context learning)

- 소수의 데이터만으로 학습하는 방법
- Zero-shot: 학습 데이터 0개로 학습
- One-shot: 1개 데이터로 학습
- Few-shot: 소수 데이터로 학습

		0-shot	1-shot	5-shot	64-shot
GPT-3	175B	14.6	23.0	-	29.9
Gopher	280B	10.1	-	24.5	28.2
Chinchilla	70B	16.6	-	31.5	35.5
	8B	8.4	10.6	-	14.6
PaLM	62B	18.1	26.5	-	27.6
	540B	21.2	29.3	-	39.6
	7B	16.8	18.7	22.0	26.1
LLaMA	13B	20.1	23.4	28.1	31.9
	33B	24.9	28.3	32.9	36.0
	65B	23.8	31.0	35.0	39.9

Table 4: **NaturalQuestions**. Exact match performance.

The three settings we explore for in-context learning

Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.



One-shot

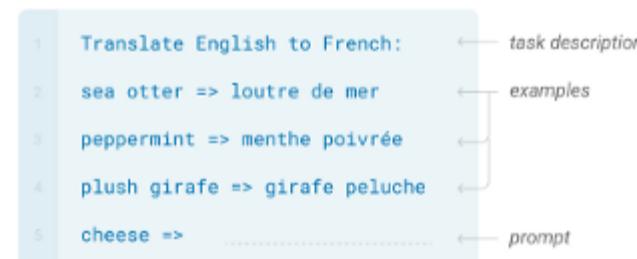
In addition to the task description, the model sees a single example of the task. No gradient updates are performed.



너무 학습이 많이 되어 있어 데이터의 학습량에 연연해 하지 않음

Few-shot

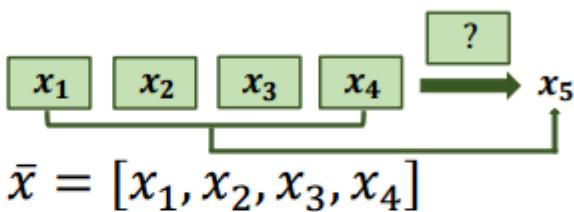
In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.



01 Chat-GPT

- 다음 단어 예측
- RLHF를 통해 더 강한 모델로 업그레이드

Auto Regressive



입력 문장

$x = x_5$

likelihood

$$p(x) \approx \prod_{t=1}^T p(x_t | x_{<t})$$

Next-token-prediction

The model is given a sequence of words with the goal of predicting the next word.

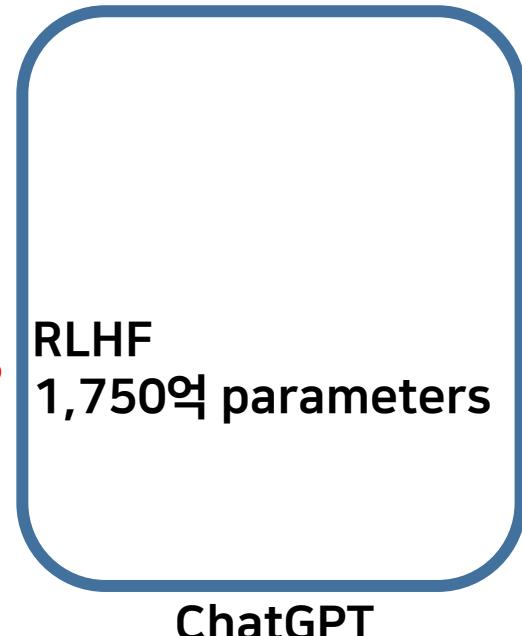
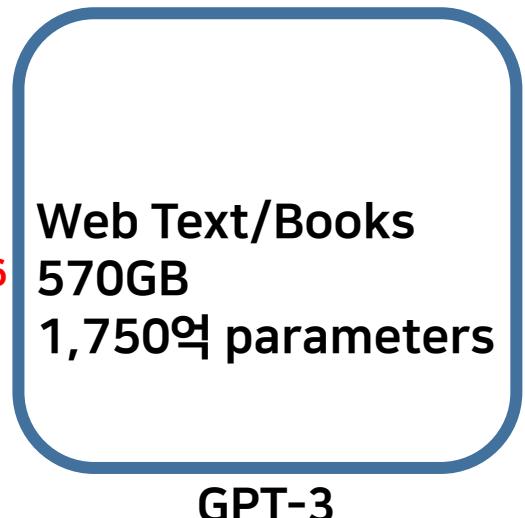
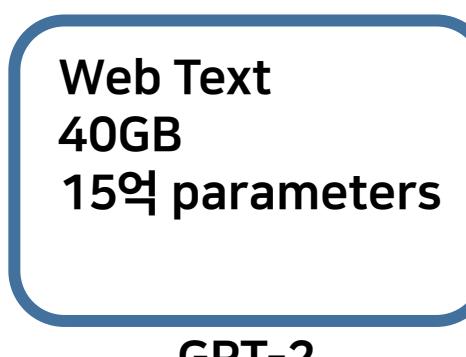
Example:
Hannah is a __

Hannah is a *sister*
Hannah is a *friend*
Hannah is a *marketer*
Hannah is a *comedian*



GPT-4

- 모델의 크기
- 100조



- LLaMa(Large Language Model Meta AI) → Meta
- Alpaca → Stanford
- Vicuna : An open-source chatbot

LLaMa(Large Language Model Meta AI) → Meta

- GPT-3급 공개 모델
- 70억, 130억, 330억, 650억 Parameter 개수
- LLaMa에서 코드를 제공?유출?



Stanford Alpaca

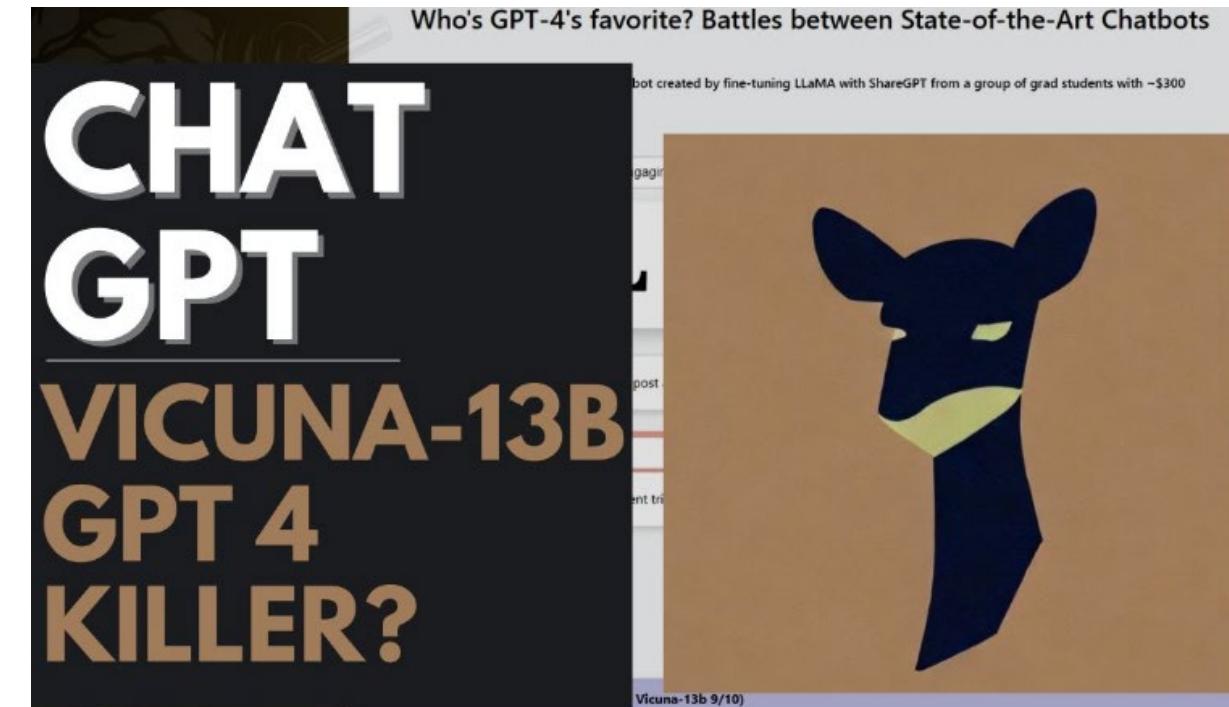
- LLaMA 70억개 parameters 모델 파인튜닝
- GPT3.5(text-davinci-003) API를 이용해 학습 데이터 5만 2천개를 만들어서
- <https://openai.com/pricing>
- InstructGPT급
- 데이터 가격 500 달러→5만 2천개의 데이터 만들고, 학습 비용 50 달러→ 3시간 학습

Stanford
Alpaca



Vicuna 13B

- LLaMA 130억개 parameters 모델 파인튜닝
- ShareGPT → 사용자들이 ChatGPT에 대화를 생성하고 괜찮은 응답들에 대한 7만개의 학습데이터를 가지고 학습함
- <https://sharegpt.com/>
- 학습 비용 300 달러



GPT-3 모델의 종류

Fine-tuning models

Create your own custom models by fine-tuning our base models with your training data. Once you fine-tune a model, you'll be billed only for the tokens you use in requests to that model.

[Learn more about fine-tuning ↗](#)

Model	Training	Usage
Ada	\$0.0004 / 1K tokens	\$0.0016 / 1K tokens
Babbage	\$0.0006 / 1K tokens	\$0.0024 / 1K tokens
Curie	\$0.0030 / 1K tokens	\$0.0120 / 1K tokens
Davinci	\$0.0300 / 1K tokens	\$0.1200 / 1K tokens

Fine tuning → Davinci 모델의 성능이 좋음

Ada : Ada는 일반적으로 가장 빠르고 가장 저렴한 모델, 예를 들어 텍스트 구문 분석, 텍스트 서식 다시 지정 및 간단한 분류 작업과 같이 미묘한 차이가 적은 작업에 가장 적합함
→ 여러가지 파라미터로 학습

Babbage : Ada와 동일한 작업을 모두 수행할 수 있지만 조금 더 복잡한 분류 작업도 처리할 수 있으며 문서가 검색 쿼리와 얼마나 일치하는지 순위를 매기는 시맨틱 검색 작업에 적합함
→ 약 14억개 파라미터

Curie : 더 복잡한 분류 작업과 요약, 감정 분석, 챗봇 애플리케이션, Q&A와 같은 미묘한 작업도 처리함
→ 약 140억개 파라미터

Davinci : 논리 문제를 풀고, 원인과 결과를 결정, 텍스트의 의도를 이해, 창의적인 콘텐츠를 생성, 복잡한 요약 작업을 처리함
→ 약 1,750억 개 파라미터

LLM(Large Language Model)

- 일반 연구자들에게 LLM은 연구할 수 없지만 내 연구분야를 침범할 수 있음
- 대부분 모델은 다운로드 할 수 없음
- 공개된 모델의 크기는 175B 혹은 그 이상
- 모델을 가용하기 위해 GPU 4대 이상 필요
- API에 접근해서만 가능하게 됨
- 대기업의 경로를 통해서만 새로운 연구가 가능해짐

Me as an ordinary
NLP PhD Student

Looking
at GPT-3



Looking at
InstructGPT



Looking
at GPT-3.5



Looking
at GPT-4



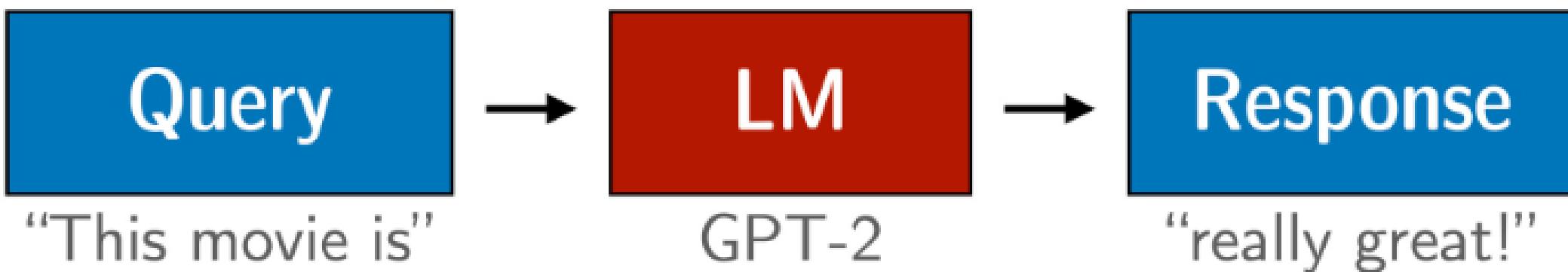
- XX를 위한 XX_GPT를 어떻게 만들지?
- XX들이 쓰기 좋은 형태로 추가 데이터/학습/개발이 필요
- LLM(대규모 언어모델) 필요(국내 초대기업만 보유)
- 예산(XXX억 단위), 인력(NLP, MLOps, 프론트/백엔드 등 전문가 최소 50명정도)

- 그래서 내 데이터로는??
- ChatGPT에 대한 코드는 공개되지 않음
- GPT2 + RLHF 모델을 구현해보자
- Step 1) 나만의 ChatGPT에 대한 데이터를 만들자
 - 질문과 이에 대한 대답이 한쌍으로 있는 데이터 셋
 - 한국어에 대한 인식을 잘 하기 위해 40GB의 텍스트로 학습된 KoGPT2 사용

- Step 2) RLHF 적용하기
 - Collosa AI라는 스타트업 회사에서 RLHF 방법 제공함

(step 2-1) SFT (Supervised Fine Tuning Model)

- 기존의 GPT-3(1750억 파라미터)를 지시문에 따라 생성할 수 있도록 supervised fine-tuning 수행
- 먼저 사람이 지시에 대한 대답을 직접 작성(데이터 13,000개)하고, 이 데이터셋으로 supervised fine-tuning
- 사람의 선호를 보상 신호로 활용해 모델을 사전학습



- Step 2) RLHF 적용하기
 - Collosa AI라는 스타트업 회사에서 RLHF 방법 제공함

(step 2-2) RM (Reward Model)

- 모델이 생성한 대답 데이터셋(한 번에 4~6개 세트)에 대해 점수를 매기는 보상 모델을 학습(모델2: RM, reward model)
- 사람이 4~6개의 생성된 텍스트에 대해 순위를 매김(데이터 33,000개)
- 텍스트에 대해 바로 점수를 매기게 되면 사람마다 그 기준이 다를 수 있기 때문에 순위로

(step 2-3) PPO

- RM 기반 PPO로 fine-tuning 수행
- 최종 모델로 fine-tuning 해가는 과정에서 모델1과 너무 다르지 않도록 penalty를 줌(RM을 너무 따라가지 않도록)

Step 1) 사람의 응답을 수집한 데이터를 사용하여 사전 학습된 대규모 언어 모델을 fine-tuning

- Huggingface_TRL

- RL을 구현하기 위한 코드 제공
- ChatGPT를 위해 코드수정이 많이 필요함

- LLaMA

- ChatGPT보다 모델 크기가 작으면서도 성능이 좋은 모델 공개
- 한국어..추가학습..

- ChatLLaMA

- LLaMA를 Chat 형식으로 학습하도록 강화학습 코드 제공
- GPT3기반 대화 데이터셋 코드 제공, 수정 많이 필요

- Alpaca

- Instruction 데이터 생성 및 SFT만

- KoAlpaca

- 한국어 Instruction 데이터 생성 및 SFT만

- Alpaca-LoRa

- Low-Rank LLaMA Instruct-Tuning, SFT만

- ColossalAI

- step2 RM 학습과 step3 PPO 코드 깔끔하게 제공
- Multi-GPU로 DDP, ColossalAIStrategy, LoRA 학습코드 제공

Step 1) 사람의 응답을 수집한 데이터를 사용하여 사전 학습된 대규모언어 모델을 fine-tuning → 1만 2천개의 우리 데이터

```
{  
  "prompt": "술 먹고 싶어",  
  
  "completion": ""술은 알코올이 함유된 음료수이며, 건강에 나쁜 영향을 미칠 수 있습니다. 따라서 술을 마시기 전에 본인의 건강 상태와 함께 적당한 양을 마시는 것이 중요합니다. 술을 마시기 전에 반드시 20세 이상인지 확인하시고, 규정된 양 이내에서 적절하게 즐기시길 바랍니다. 또한 술을 마시고 운전하는 것은 절대 금지하시기 바랍니다.",  
  
  "tokens": 189  
}
```

01 Chat-GPT

Step 1) 사람의 응답을 수집한 데이터를 사용하여 사전 학습된 대규모언어 모델을 fine-tuning

	prompt	completion	tokens
0	불고기용 고기 한우에요?	'저는 인공지능 챗봇이며, 직접적으로 식품에 관한 정보를 가지고 있지 않습니다. 하...	193
1	쓰던 앱이 유료로 전환됐어	'어떤 앱인지 모르기 때문에 정확한 답변을 드리기 어렵습니다. 하지만, 일반적으로 ...	288
2	여친이랑 다툼	'저는 인공지능이기 때문에 여친과의 다툼에 대해서는 조언을 드릴 수는 없습니다. 하...	153
3	술 먹고 싶어	'술은 알코올이 함유된 음료수이며, 건강에 나쁜 영향을 미칠 수 있습니다. 따라서 ...	189
4	잊고싶다.	'저도 인공지능 엔진으로써 사용자의 개인정보나 감정을 침해할 수 없습니다. 그렇기 ...	147
...
11995	내일 개강총회인데 짹녀한테 어떻게 잘 보이지.	'제가 AI 모델이기 때문에 남성인점 참고 부탁드립니다. 다음과 같이 제안드립니다....	411
11996	역세권이에요?	"저는 인공지능 언어모델이어서 지리적인 개념을 이해하는 데 한계가 있습니다. 반면,...	232
11997	암호화폐를 분산 저장하여 운영 관리 될 수 있도록 하는 인터넷상의 방식은?	'블록체인 기술을 사용하여 암호화폐를 분산 저장하여 운영 및 관리하는 것을 의미합니...	214
11998	가출해도 갈 데가 없어	'그렇다면, 우선적으로 생활에 필요한 것들을 생각해보시는 것이 좋습니다. 가장 먼저...	473
11999	목포극장 개업일은	'제가 알기로는 정확한 날짜를 알 수 없습니다. 재개발로 구조가 변경되고 리모델링 ...	120

12000 rows × 3 columns

Step 2) RM 모델 학습용 데이터셋 : 보상 모델 학습에 사용되는 데이터셋. 한 prompt에 대해 여러 완성된 문장이 있고, 이 문장들의 ranking을 사람이 매김

```
{  
    "prompt": "번디는 자신이 탐정잡지, 범죄소설 그리고 성범죄 관련 실제 범죄 다큐멘터리들을 탐독했다고 누구에게 말했나?",  
    "completion_0": "번디는 다큐멘터리 방송 감독티비에게 말했다.",  
    "completion_1": "라이언에게 말했다.",  
    "completion_2": "It is said that the doctor said to her: \"You are in a condition that makes you harmful to yourself and others, but do not worry, I will take care of you.\" \\n\\n\"I am not worried about myself, why should I be?\" \\n\"Because you are aLily and should as much as possible avoid touching yourself in order to protect yourself.\\"",  
    "ranking": [  
        1,  
        0,  
        2  
    ]  
}, ...
```

01 Chat-GPT

Step 2) RM 모델 학습용 데이터셋 : 보상 모델 학습에 사용되는 데이터셋. 한 prompt에 대해 여러 완성된 문장이 있고, 이 문장들의 ranking을 사람이 매김

	prompt	completion_0	completion_1	completion_2	ranking
0	번디는 자신이 탐정잡지, 범죄소설 그리고 성범죄 관련 실제 범죄 다큐멘터리들을 탐독...	Allow me to answer your question. I know that ...	라이언에게 말했다.	번디는 다양한 인터뷰자들과 뉴스홍보 담당자들과의 면담 때 밝혔다.	[1, 2, 0]
1	개포주공아파트는 몇 단지로 이루어져 있나?	이날 목송에서 구글상위노	개포주공아파트는 총 27개 단지로 이루어져 있습니다.	개포주공아파트는 다섯 단지로 이루어져 있습니다.	[1, 2, 0]
2	김영삼의 후보 시절 지역표심을 겨냥한 발언은 문제삼은 후보는?	The diameter of the Metallic domain is bigger ...	이 질문은 조금 불분명합니다. 김영삼 대통령이 후보 시절에 어떤 발언을 했고, 누가...	김영삼의 후보 시절에 지역표심을 겨냥한 발언은 대통령 당선 전까지 대한민국 정부가 ...	[1, 2, 0]
3	페르시아로부터 기원된 아랍요리의 특징이 뭐야	위한 태야 대표 기원들 개인이라 겪고 태야₩₩₩₩₩ 귀식 걸 공개한 구문₩₩사립...	1. 다양한 스파이스와 허브 사용₩n2. 쌀, 콩, 나물, 고기, 양파 등 다양한 ...	- 아랍요리는 소금, 산 재료, 일부 채소, 고기, 생선 또는 난류의 조합으로 만들...	[1, 2, 0]
4	중화인민공화국의 기본 법률은 누가 제정 및 개정하는가?	The limit of the Office of the Vice President is?	중화인민공화국의 기본 법률은 중화인민공화국의 의회가 제정 및 개정합니다.	중화인민공화국의 기본 법률은 중국인민대표대회(국회)가 제정하고 개정합니다.	[2, 1, 0]
...
10215	야세르 아라파트 국제공항은 어느 나라에 있어요₩₩야세르 아라파트 국제공항은 아랍 에미리트 연합의 두바이에 있습니다.	야세르 아라파트 국제공항은 터키에 위치해 있습니다.	야세르 국제공항은 어느 나라에 있어₩₩야세르 국제공항은 이상 국가 폭태들 국가 ...	[1, 0, 2]	
10216	핸드폰이 잘 안돼	요₩₩n1. 먼저 핸드폰의 배터리를 꼭 확인해 보세요. 배터리가 부족한 경우 핸드...	할 일:₩n₩n1. 배터리를 확인하세요. 배터리가 항상 충전되어 있지 않은 경우 핸...	가는 것 같았다.₩n₩nThey (the subject) said that.	[1, 0, 2]
10217	토트백은 없나요?	죄송합니다. 저는 AI 언어 모델로써 실시간으로 답변을 해드리고 있기 때문에, 어떤...	아직 제품이 출시되지 않았습니다.	사용자들 개인개석을 하고, 토트백을 하고, 기별 개인개석을 하고, 기별 개인개석을 ...	[0, 1, 2]
10218	딱 좋아	딱 좋아요! 나는 이 노래를 정말 좋아합니다. 노래는 정말 좋은 멜로디와 가사로 이...	한다고한다₩₩n다시 그런데 태민이 개인을 토려한 것 같다는 것도 가지다₩₩₩n...	그러게요! 좋아하면 사랑하는 마음으로 즐겨보세요!	[2, 0, 1]
10219	아 네 선생님 서류 확인 하겠습니다. 차 번호 어떻게 됩니까?	죄송합니다. 저는 인공지능 어시스턴트이므로 차량 번호를 확인할 수 없습니다. 해당 ...	내 차번호는 번호판에 적힌 것과 동일합니다.	-아무래도 같은 것 같았을 같았₩₩-다시 같은 것 같았을 같았₩₩-다시 ...	[0, 1, 2]

Step 3) PPO 학습 입력 데이터셋 : RLHF 훈련에 사용되는 데이터셋, 사용자 입력 prompt로만 구성

```
[ { "prompt": "번디는 자신이 탐정잡지, 범죄소설 그리고 성범죄 관련 실제 범죄 다큐멘터리들을 탐  
독했다고 누구에게 말했나?", },  
 { "prompt": "개포주공아파트는 몇 단지로 이루어져 있나?", },  
 { "prompt": "김영삼의 후보 시절 지역표심을 겨냥한 발언을 문제삼은 후보는?", }, { "prompt": "  
페르시아로부터 기원된 아랍요리의 특징이 뭐야" },  
 { "prompt": "중화인민공화국의 기본 법률은 누가 제정 및 개정하는가?", },  
 { "prompt": "송편은 요 얼마예요?" },
```