23-1 비정형데이터분석

GPT series and Language Models

Youtube 요약 영상 과제 3

고려대학교 산업경영공학과

팀 : 지란지교

팀원 : 박새란 이지윤

Content

GPT Series

1. GPT-1
2. GPT-2
3. GPT-3
4. GPT-3 Family
5. GPT-4

Language Models

6. MT-DNN
7. MASS
8. UNILM
9. XLNet
10. RoBERTa
11. ALBERT

GPT Series

✓ Overview

- GPT-1 : generative pre-training을 통해 transfer task에 대해 효과적인 성능을 보임.
- GPT-2 : zero-shot setting, 다양한 태스크에 좋은 성능을 찍음
- GPT-3 : 더 많은 데이터셋을 이용한 사전 학습. Fine-tuning 없이 few-shot setting에서의 향상된 성능 증명
- GPT-3.5 : text + code데이터를 활용
    - Instruct-GPT : 날 것 그대로의 인터넷 데이터으로 인한 생성된 텍스트의 문제. -> human feedback을 통한 강화학습
    - Codex : code 데이터 포함
    - ChatGPT : 안정성을 고려한 모델링 (RLHF)
- GPT-4
    - Multi-modal
    - Hallucination 보완
    - 더 어려운 문제 해결 가능

01

GPT-1

GPT-1 : improving language understanding by generative pre-training

✓ Objective and Concepts

- 문제점 : 수많은 NLU task와 부족한 labeled dataset
- 해결 :수많은 unlabeled data 활용 ➡️ *"Generative Pre-training"*, Unsupervised Learning
- 결과 : 12개 task 중 9개에서 SOTA 달성

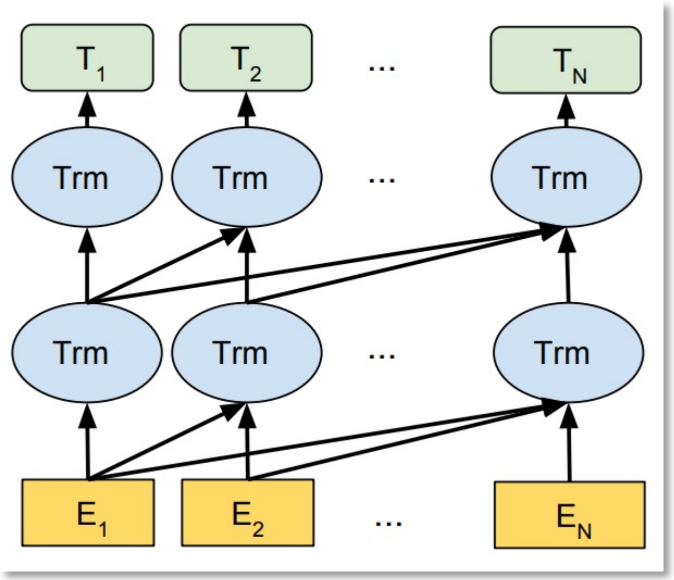
Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

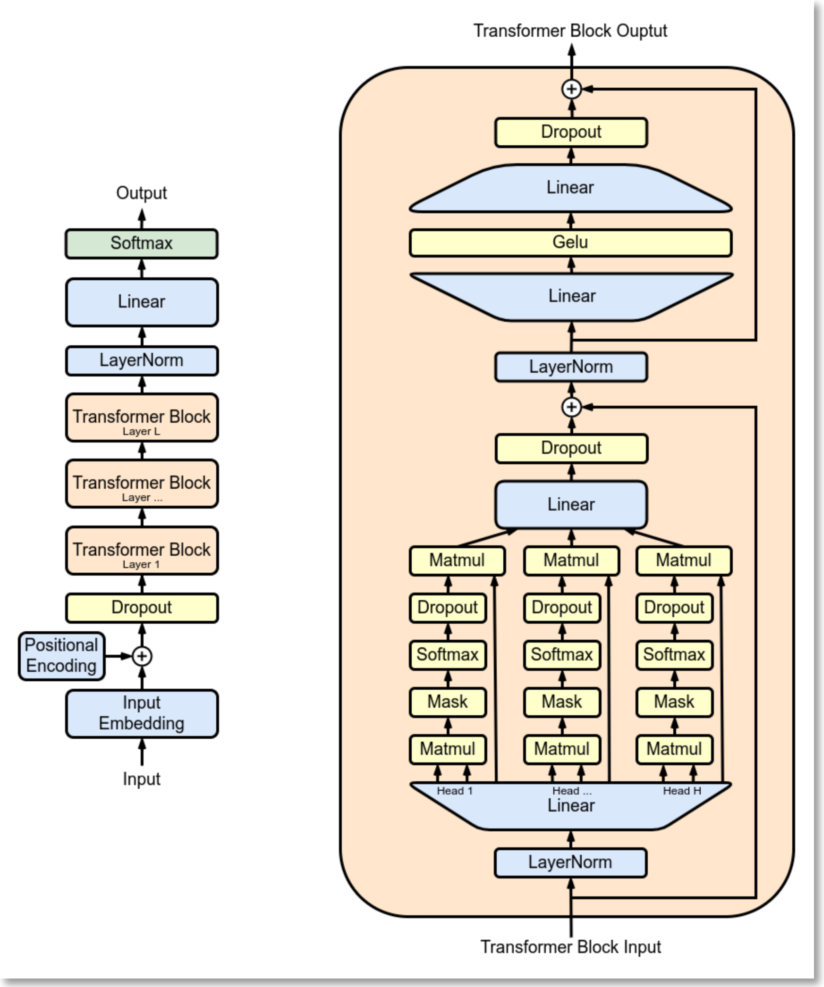
GPT-1 : improving language understanding by generative pre-training

✓ Objective and Concepts

✓ Decoder로만 이루어진 모델

- Masked self-attention





https://en.wikipedia.org/wiki/Generative_pre-trained_transformer

GPT-1 : improving language understanding by generative pre-training

✓ **Objective and Concepts**

✓ Unsupervised Learning

Given an unsupervised corpus of tokens $\mathcal{U} = \{u_1, \ldots, u_n\}$, we use a standard language modeling objective to maximize the following likelihood:

$$L_1(\mathcal{U}) = \sum_i \log P(u_i | u_{i-k}, \ldots, u_{i-1}; \Theta) \qquad (1)$$

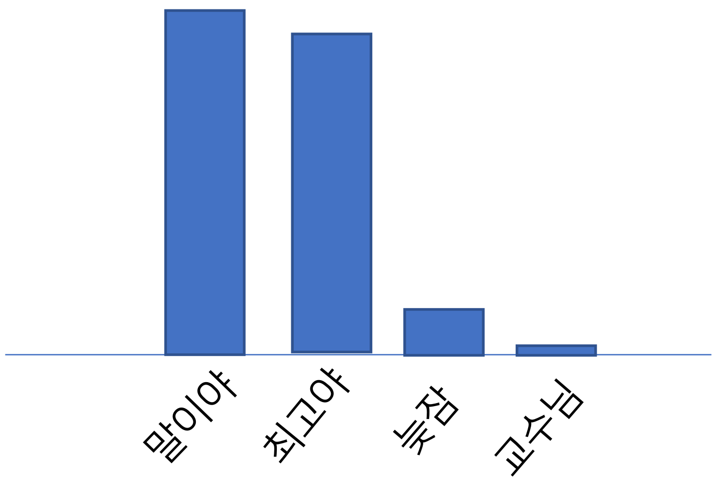
Q : 연세, 꼬마야, 왜, 자꾸, 틀리는, 거야, 고연전이란, < ? >

A : logP(말이야 | 연세, 꼬마야, 왜, 자꾸, 틀리는, 거야, 고연전이란)

A : log P(최고야 | 연세, 꼬마야, 왜, 자꾸, 틀리는, 거야, 고연전이란)

A : log P(늦잠 | 연세, 꼬마야, 왜, 자꾸, 틀리는, 거야, 고연전이란)

A : log P(교수님 | 연세, 꼬마야, 왜, 자꾸, 틀리는, 거야, 고연전이란)



Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : improving language understanding by generative pre-training

✓ Objective and Concepts

✓ Generative Pre-training : Unsupervised Learning

✓ Supervised Learning

- Task-aware input : Figure1
- Textual entailment
    - "<s> + premise + $ + hypothsis + <e>"
- Similarity
    - "<s> + Sen1 + $ + Sen2 + <e>"
    - "<s> + Sen2 + $ + Sen1 + <e>"
    - 두 문장의 순서에 대한 정보가 필요없기 때문
- QA
    - "<s> + document + question+ $ + answer + <e>"
    - context : document + question

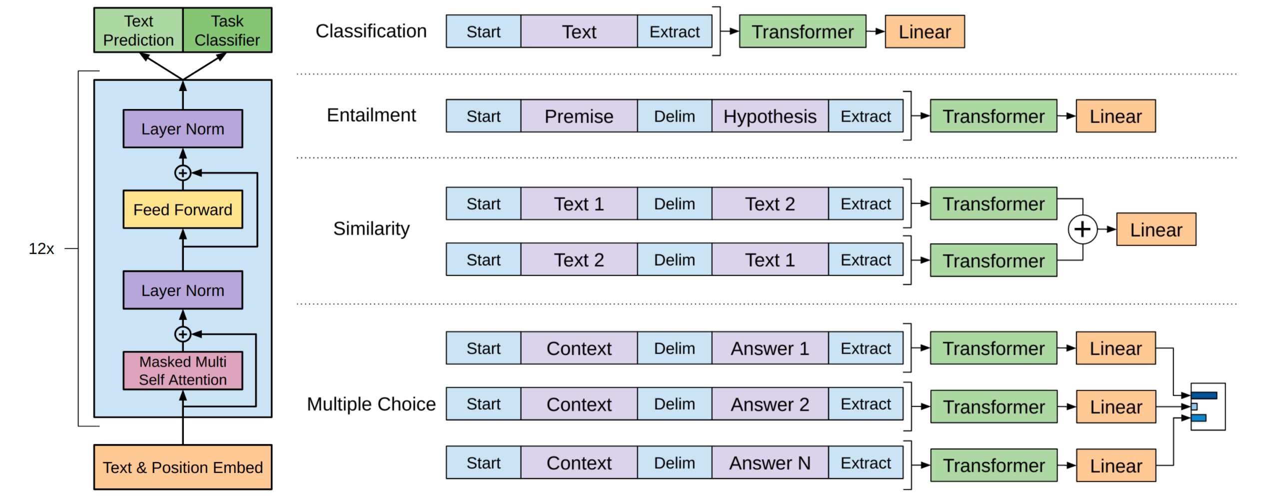


Figure 1: **(left)** Transformer architecture and training objectives used in this work. **(right)** Input transformations for fine-tuning on different tasks. We convert all structured inputs into token sequences to be processed by our pre-trained model, followed by a linear+softmax layer.

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : improving language understanding by generative pre-training

✓ Objective and Concepts

✓ Generative Pre-training : Unsupervised Learning

✓ Supervised Learning

- Auxiliary learning objective : language modeling
  - Lambda를 통해 조절
  - 효과 : better generalization, faster convergence

Task에 대한 정답

$$L_2(\mathcal{C}) = \sum_{(x,y)} \log P(\boxed{y}|x^1, \ldots, x^m).$$

$$L_3(\mathcal{C}) = L_2(\mathcal{C}) + \lambda L_1(\mathcal{C})$$

Language modeling objective

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : improving language understanding by generative pre-training

✓ Dataset

- Pre-training : BookCorpus
  - 7000개의 다양한 장르의 책
- Supervised Learning

Table 1: A list of the different tasks and datasets used in our experiments.

Task	Datasets
Natural language inference	SNLI [5], MultiNLI [66], Question NLI [64], RTE [4], SciTail [25]
Question Answering	RACE [30], Story Cloze [40]
Sentence similarity	MSR Paraphrase Corpus [14], Quora Question Pairs [9], STS Benchmark [6]
Classification	Stanford Sentiment Treebank-2 [54], CoLA [65]

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : improving language understanding by generative pre-training

✓ Performance

✓ Natural Language Inference

- *Given the strong performance of our approach on larger NLI datasets, it is likely our model will benefit from multi-task training as well but we have not explored this currently.* ➡️ GPT-2 논문의 제목

Table 2: Experimental results on natural language inference tasks, comparing our model with current state-of-the-art methods. 5x indicates an ensemble of 5 models. All datasets use accuracy as the evaluation metric.

Method	MNLI-m	MNLI-mm	SNLI	SciTail	QNLI	RTE
ESIM + ELMo [44] (5x)	-	-	89.3	-	-	-
CAFE [58] (5x)	80.2	79.0	89.3	-	-	-
Stochastic Answer Network [35] (3x)	80.6	80.1	-	-	-	-
CAFE [58]	78.7	77.9	88.5	83.3		
GenSen [64]	71.4	71.3	-	-	82.3	59.2
Multi-task BiLSTM + Attn [64]	72.2	72.1	-	-	82.1	**61.7**
Finetuned Transformer LM (ours)	**82.1**	**81.4**	**89.9**	**88.3**	**88.1**	56.0

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : improving language understanding by generative pre-training

✓ **Performance**

✓ QA & commonsense reasoning

- Story Cloze, RACE를 통해 긴 context에 대해서도 다룰 수 있다는 걸 증명함.

Table 3: Results on question answering and commonsense reasoning, comparing our model w current state-of-the-art methods.. 9x means an ensemble of 9 models.

Method	Story Cloze	RACE-m	RACE-h	RACE
val-LS-skip [55]	76.5	-	-	-
Hidden Coherence Model [7]	77.6	-	-	-
Dynamic Fusion Net [67] (9x)	-	55.6	49.4	51.2
BiAttention MRU [59] (9x)	-	60.2	50.3	53.3
Finetuned Transformer LM (ours)	**86.5**	**62.9**	**57.4**	**59.0**

Passage:
In a small village in England about 150 years ago, a mail coach was standing on the street. It didn't come to that village often. People had to pay a lot to get a letter. The person who sent the letter didn't have to pay the postage, while the receiver had to.
"Here's a letter for Miss Alice Brown," said the mailman.
" I'm Alice Brown," a girl of about 18 said in a low voice.
Alice looked at the envelope for a minute, and then handed it back to the mailman.
"I'm sorry I can't take it, I don't have enough money to pay it", she said.
A gentleman standing around were very sorry for her. Then he came up and paid the postage for her.
When the gentleman gave the letter to her, she said with a smile, " Thank you very much, This letter is from Tom. I'm going to marry him. He went to London to look for work. I've waited a long time for this letter, but now I don't need it, there is nothing in it."
"Really? How do you know that?" the gentleman said in surprise.
"He told me that he would put some signs on the envelope. Look, sir, this cross in the corner means that he is well and this circle means he has found work. That's good news."
The gentleman was Sir Rowland Hill. He didn't forgot Alice and her letter.
"The postage to be paid by the receiver has to be changed," he said to himself and had a good plan.
"The postage has to be much lower, what about a penny? And the person who sends the letter pays the postage. He has to buy a stamp and put it on the envelope." he said . The government accepted his plan. Then the first stamp was put out in 1840. It was called the "Penny Black". It had a picture of the Queen on it.

Questions:

1): The first postage stamp was made _.
A. in England B. in America C. by Alice D. in 1910

2): The girl handed the letter back to the mailman because _ .
A. she didn't know whose letter it was
B. she had no money to pay the postage
C. she received the letter but she didn't want to open it
D. she had already known what was written in the letter

3): We can know from Alice's words that _ .
A. Tom had told her what the signs meant before leaving
B. Alice was clever and could guess the meaning of the signs
C. Alice had put the signs on the envelope herself
D. Tom had put the signs as Alice had told him to

4): The idea of using stamps was thought of by _ .
A. the government
B. Sir Rowland Hill
C. Alice Brown
D. Tom

5): From the passage we know the high postage made _ .
A. people never send each other letters
B. lovers almost lose every touch with each other
C. people try their best to avoid paying it
D. receivers refuse to pay the coming letters

Answer: ADABC

Table 1: Sample reading comprehension problems from our dataset.

RACE: Large-scale ReAding Comprehension Dataset From Examinations. Lai et al.
Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : improving language understanding by generative pre-training

- ✓ **Performance**

- ✓ Semantic Similarity and Classification

Table 4: Semantic similarity and classification results, comparing our model with current state-of-the-art methods. All task evaluations in this table were done using the GLUE benchmark. (*mc*= Mathews correlation, *acc*=Accuracy, *pc*=Pearson correlation)

Method	**Classification**		**Semantic Similarity**			**GLUE**
	CoLA (mc)	SST2 (acc)	MRPC (F1)	STSB (pc)	QQP (F1)	
Sparse byte mLSTM [16]	-	**93.2**	-	-	-	-
TF-KLD [23]	-	-	**86.0**	-	-	-
ECNU (mixed ensemble) [60]	-	-	-	81.0	-	-
Single-task BiLSTM + ELMo + Attn [64]	35.0	90.2	80.2	55.5	66.1	64.8
Multi-task BiLSTM + ELMo + Attn [64]	18.9	91.6	83.5	72.8	63.3	68.9
Finetuned Transformer LM (ours)	**45.4**	91.3	82.3	**82.0**	**70.3**	**72.8**

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

GPT-1 : Improving language understanding by generative pre-training

✓ Performance

- Layer가 수가 많아질수록 향상되는 transfer 능력
- 사전 학습 횟수가 많아질수록 zero-shot setting에서 성능이 향상됨.

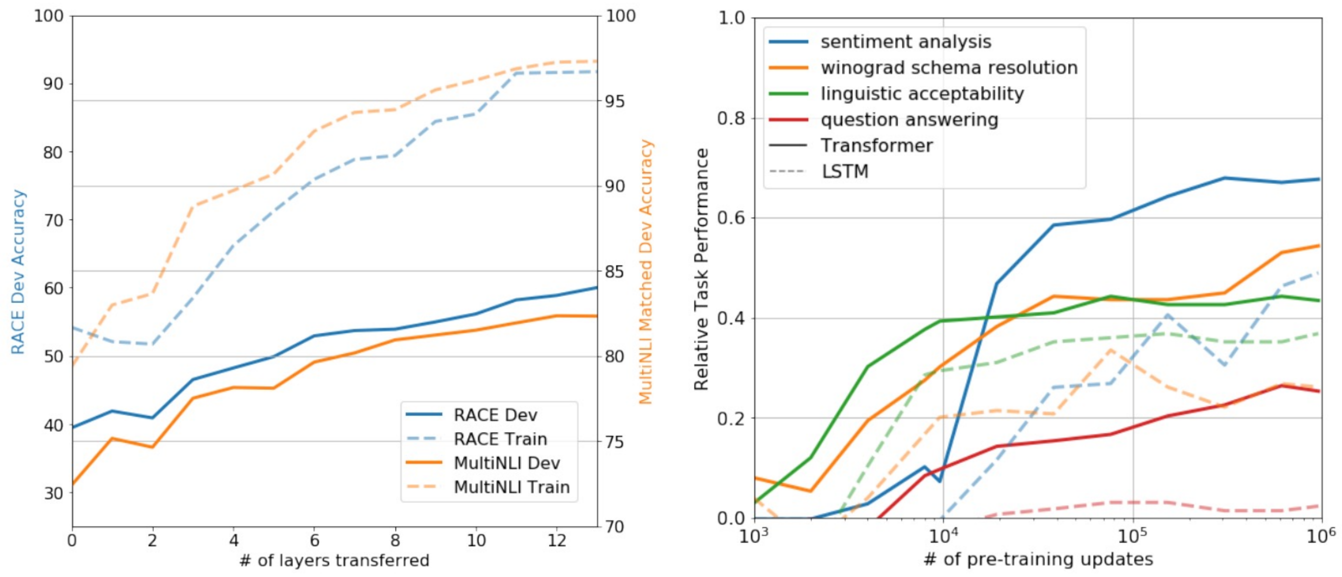


Figure 2: (**left**) Effect of transferring increasing number of layers from the pre-trained language model on RACE and MultiNLI. (**right**) Plot showing the evolution of zero-shot performance on different tasks as a function of LM pre-training updates. Performance per task is normalized between a random guess baseline and the current state-of-the-art with a single model.

Radford, Alec, et al. "Improving language understanding by generative pre-training." (2018).

02

GPT-2

GPT-2 : Language models are unsupervised multitask learners

✓ Objective and Concepts

- Supervision 없이 다양한 task를 해결 할 수 있는 GPT-2
- Document, Question ➡️ 🤖 ➡️ Answer
- 결과 : 15억개의 파라미터를 가진 모델, GPT-2는 zero-shot setting에서 8개의 task 중 7개에서 SOTA 달성

GPT-2 : Language models are unsupervised multitask learners

✓ **Concepts**

✓ Zero-shot Setting

- 모든 태스크에 대해서 fine-tuning 학에는 태스크에 대한 데이터가 없거나 부족함.
- Language modeling 사전학습이라는 특징과 비슷하게, Task description에 대해 conditioning을 주고 텍스트를 생성해보자.
- 효과 : unsupervised objective에서 global minimum이 결국 supervised objective에서 global minimum이 된다.
  - *Unsupervised Learning과 Supervised Learning에 대한 Discrepancy 걱정을 할 필요 없음*

- (Task Description, Query, Answer)
- (Answer the Question, document, question, answer)
- (translate to french, english text, french text)

Radford, Alec, et al. "Language models are unsupervised multitask learners." OpenAI blog 1.8 (2019): 9.

GPT-2 : Language models are unsupervised multitask learners

✓ Dataset

- 문제점 : GPT-1에서는 1개의 domain에 치우쳐진 BookCorpus 데이터셋으로만 학습함.
- 제안 : 다양한 domain과 context를 가진 데이터셋으로 학습 시키자.
  - WebText : 40GB, 800만개 문서
  - Reddit에서 크롤한 후 투표를 통해 데이터로 결정함
  - 이미 task와 관련된 text가 많이 포함되어 있음 (Table1)

”I’m not the cleverest man in the world, but like they say in French: **Je ne suis pas un imbecile [I’m not a fool].**

In a now-deleted post from Aug. 16, Soheil Eid, Tory candidate in the riding of Joliette, wrote in French: ”**Mentez mentez, il en restera toujours quelque chose**,” which translates as, ”**Lie lie and something will always remain.**”

“I hate the word ‘**perfume**,”’ Burr says. ‘It’s somewhat better in French: ‘**parfum**.’

If listened carefully at 29:55, a conversation can be heard between two guys in French: “**-Comment on fait pour aller de l’autre coté? -Quel autre coté?**”, which means “**- How do you get to the other side? - What side?**”.

If this sounds like a bit of a stretch, consider this question in French: **As-tu aller au cinéma?**, or **Did you go to the movies?**, which literally translates as Have-you to go to movies/theater?

“**Brevet Sans Garantie Du Gouvernement**”, translated to English: “**Patented without government warranty**”.

Table 1. Examples of naturally occurring demonstrations of English to French and French to English translation found throughout the WebText training set.

Radford, Alec, et al. "Language models are unsupervised multitask learners." OpenAI blog 1.8 (2019): 9.

GPT-2 : Language models are unsupervised multitask learners

✓ Performance

- Zero-shot Setting에서 좋은 성능을 보임.
  - Long-range dependency가 있는 텍스트에 대해서도 잘 함.
  - MRC task 잘 함.
  - Translation 잘 함.
  - Summarization은 잘 못함.

	LAMBADA (PPL)	LAMBADA (ACC)	CBT-CN (ACC)	CBT-NE (ACC)	WikiText2 (PPL)	PTB (PPL)	enwik8 (BPB)	text8 (BPC)	WikiText103 (PPL)	1BW (PPL)
SOTA	99.8	59.23	85.7	82.3	39.14	46.54	0.99	1.08	18.3	**21.8**
117M	**35.13**	45.99	**87.65**	**83.4**	**29.41**	65.85	1.16	1.17	37.50	75.20
345M	**15.60**	55.48	**92.35**	**87.1**	**22.76**	47.33	1.01	**1.06**	26.37	55.72
762M	**10.87**	**60.12**	**93.45**	**88.0**	**19.93**	**40.31**	**0.97**	**1.02**	22.05	44.575
1542M	**8.63**	**63.24**	**93.30**	**89.05**	**18.34**	**35.76**	**0.93**	**0.98**	**17.48**	42.16

Table 3. Zero-shot results on many datasets. No training or fine-tuning was performed for any of these results. PTB and WikiText-2 results are from (Gong et al., 2018). CBT results are from (Bajgar et al., 2016). LAMBADA accuracy result is from (Hoang et al., 2018) and LAMBADA perplexity result is from (Grave et al., 2016). Other results are from (Dai et al., 2019).

Radford, Alec, et al. "Language models are unsupervised multitask learners." OpenAI blog 1.8 (2019): 9.

03

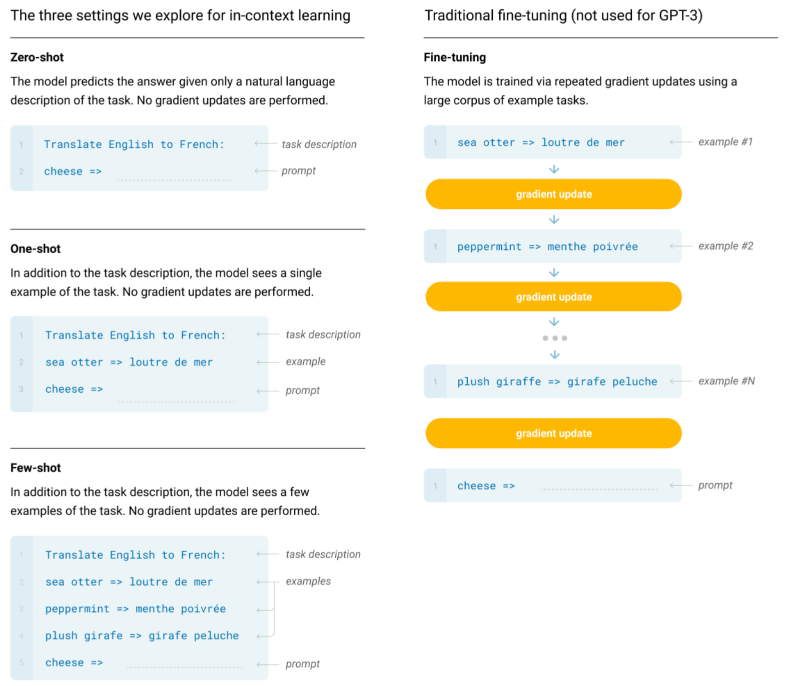
GPT-3

GPT-3 : Language model are few shot learners

✓ Concepts : In-context learning

- 얼마나 task-specific data에 의존하는가?
  - Task description : task에 대한 설명
  - Example : task와 관련된 input-output 예시
  - Prompt : 수행해야할 task에 대한 문제

methods	Task description	example	prompt
Zero-shot	O	X	O
One-shot	O	1	O
Few-shot	O	2~	O
Fine-tuning	X	N	O



Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Concepts : In-context learning

- Task : character manipulation
    - Ex) Text에서 “=“와 같은 symbol을 제거
- 결과
    - X축 examples수가 많아질수록 성능이 증가함.
    - Model 파라미터가 많을수록 성능이 좋음.
- Context가 풍부해질수록 성능이 향상됨.

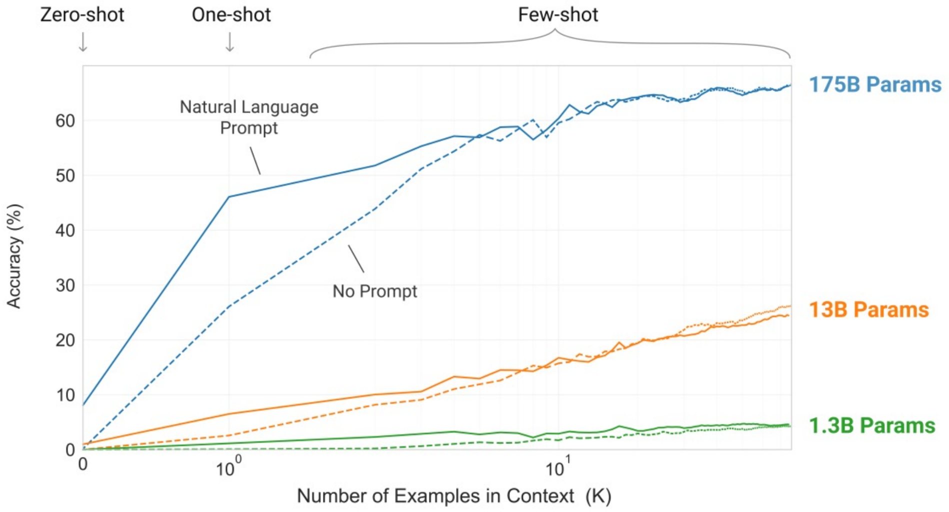


Figure 1.2: Larger models make increasingly efficient use of in-context information. We show in-context learning performance on a simple task requiring the model to remove random symbols from a word, both with and without a natural language task description (see Sec. 3.9.2). The steeper “in-context learning curves” for large models demonstrate improved ability to learn a task from contextual information. We see qualitatively similar behavior across a wide range of tasks.

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Dataset

- GPT2보다 더 많고 다양한 데이터셋

Unfortunately, a bug in the filtering caused us to ignore some overlaps, and due to the cost of training it was not feasible to retrain the model. In Section 4 we characterize the impact of the remaining overlaps, and in future work we will more aggressively remove data contamination.

Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

Table 2.2: Datasets used to train GPT-3. "Weight in training mix" refers to the fraction of examples during training that are drawn from a given dataset, which we intentionally do not make proportional to the size of the dataset. As a result, when we train for 300 billion tokens, some datasets are seen up to 3.4 times during training while other datasets are seen less than once.

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Model Architecture

- GPT2와 동일한 구조이지만 더 커진 사이즈
- Sparse attention mechanism 사용 : $O(N^2)$ ➡️ $O(N\sqrt[2]{N})$
  - Inference 속도가 훨씬 빨라짐 (약 30배)

Model Name	n_{params}	n_{layers}	d_{model}	n_{heads}	d_{head}	Batch Size	Learning Rate
GPT-3 Small	125M	12	768	12	64	0.5M	6.0×10^{-4}
GPT-3 Medium	350M	24	1024	16	64	0.5M	3.0×10^{-4}
GPT-3 Large	760M	24	1536	16	96	0.5M	2.5×10^{-4}
GPT-3 XL	1.3B	24	2048	24	128	1M	2.0×10^{-4}
GPT-3 2.7B	2.7B	32	2560	32	80	1M	1.6×10^{-4}
GPT-3 6.7B	6.7B	32	4096	32	128	2M	1.2×10^{-4}
GPT-3 13B	13.0B	40	5140	40	128	2M	1.0×10^{-4}
GPT-3 175B or "GPT-3"	175.0B	96	12288	96	128	3.2M	0.6×10^{-4}

Table 2.1: Sizes, architectures, and learning hyper-parameters (batch size in tokens and learning rate) of the models which we trained. All models were trained for a total of 300 billion tokens.

Parameters	Layers	d_{model}
117M	12	768
345M	24	1024
762M	36	1280
1542M	48	1600

Table 2. Architecture hyperparameters for the 4 model sizes.

GPT2와 비교

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Performance : LAMBADA

- long-range dependencies가 잘 모델링 되었는가?
    - long context가 주어지면, 마지막에 들어갈 단어를 예측하는 태스크
- 결과
    - Zero-shot SOTA인 Turing-NLG보다 더 좋은 성능을 보임. (76%)
    - Few-shot
        - 몇개의 단어를 주고 마지막 단어를 맞추도록 함. (86.4%)

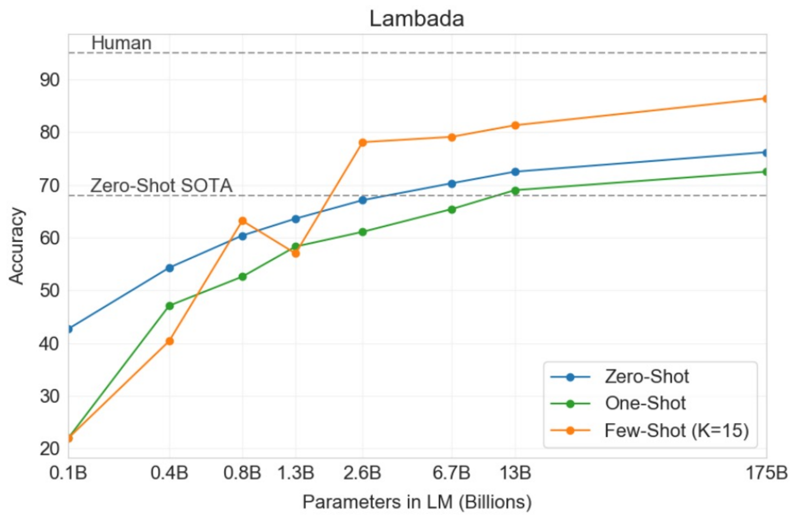


Figure 3.2: On LAMBADA, the few-shot capability of language models results in a strong boost to accuracy. GPT-3 2.7B outperforms the SOTA 17B parameter Turing-NLG [Tur20] in this setting, and GPT-3 175B advances the state of the art by 18%. Note zero-shot uses a different format from one-shot and few-shot as described in the text.

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Performance : Closed Book QA

- 넓은 factual knowledge기반으로 문제에 답변할 수 있는가 ?
- 결과
  - TriviaQA : zero-shot은 T5보다 향상됨.
  - WebQS : 0S, 1S, FS간의 gap이 큼.
    - 원인 : GPT3의 학습데이터셋의 answer style과 많이 다르기 때문임. Examples을 통해서 adaption 된 것을 볼 수 있음.
  - NQ : 다른 데이터셋에 비해 성능이 잘 나오지 않음.
    - 원인 : Wikipedia에 fine-grained knowledge가 필요한 질문이 많음.

Setting	NaturalQS	WebQS	TriviaQA
RAG (Fine-tuned, Open-Domain) [LPP+20]	**44.5**	**45.5**	**68.0**
T5-11B+SSM (Fine-tuned, Closed-Book) [RRS20]	36.6	44.7	60.5
T5-11B (Fine-tuned, Closed-Book)	34.5	37.4	50.1
GPT-3 Zero-Shot	14.6	14.4	64.3
GPT-3 One-Shot	23.0	25.3	**68.0**
GPT-3 Few-Shot	29.9	41.5	**71.2**

Table 3.3: Results on three Open-Domain QA tasks. GPT-3 is shown in the few-, one-, and zero-shot settings, as compared to prior SOTA results for closed book and open domain settings. TriviaQA few-shot result is evaluated on the wiki split test server.

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Performance : Translation

- 결과
  - 영어로 변환하는 것은 잘 되지만, 영어를 다른 언어로 변환하는 것은 부족함.
  - 프랑스어, 독일어 -> 영어로 변환하는 것은 SOTA를 찍음.
  - 영어 -> 독일어의 낮은 성능의 원인
    - GPT의 byte-level BPE tokenizer가 영어로 학습되었기 때문임.

Setting	En→Fr	Fr→En	En→De	De→En	En→Ro	Ro→En
SOTA (Supervised)	**45.6**[a]	35.0 [b]	**41.2**[c]	40.2[d]	**38.5**[e]	**39.9**[e]
XLM [LC19]	33.4	33.3	26.4	34.3	33.3	31.8
MASS [STQ+19]	37.5	34.9	28.3	35.2	35.2	33.1
mBART [LGG+20]	-	-	29.8	34.0	35.0	30.5
GPT-3 Zero-Shot	25.2	21.2	24.6	27.2	14.1	19.9
GPT-3 One-Shot	28.3	33.7	26.2	30.4	20.6	38.6
GPT-3 Few-Shot	32.6	39.2	29.7	40.6	21.0	39.5

Table 3.4: Few-shot GPT-3 outperforms previous unsupervised NMT work by 5 BLEU when translating into English reflecting its strength as an English LM. We report BLEU scores on the WMT'14 Fr↔En,

Brown, Tom, et al. "Language models are few-shot learners." *Advances in neural information processing systems* 33 (2020): 1877-1901.

GPT-3 : Language model are few shot learners

✓ Performance

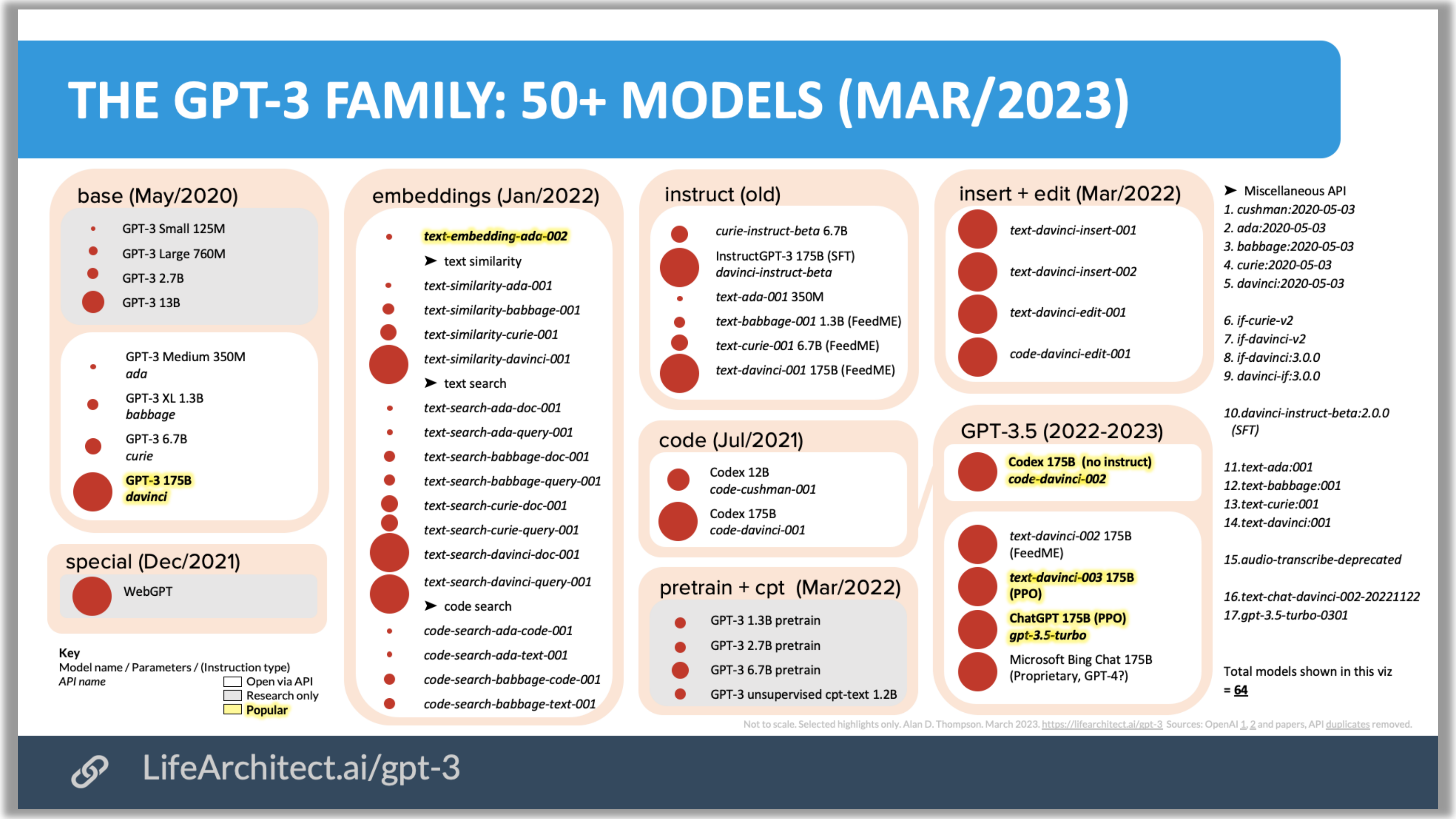
- Translation, QA, cloze task에서 성능이 좋음
- Domain adaption 능력도 좋음. (in-context learning)
- 사람만큼 글을 잘 씀. 하지만, 긴 길이의 문장을 작성하면 동일한 문장을 반복하고 내용의 일관성을 잃음.

04

GPT-3 Family

GPT-3 Family

✓ pre-trained GPT-3를 기반으로 목적에 맞게 fine-tuning하여 출시됨.



GPT-3 Family

✓ GPT-3

- fine-tuning 방식의 종류에 따라 구분
  - Davinci : 복잡한 의도, 원인과 결과, 청중을 위한 요약
  - turbo : chatGPT와 동일한 모델, 대화와 텍스트 생성 중심
  - curie : 언어번역, 감성 분류, 요약 (babbage보다는 더 복잡한 분류)
  - babbage : 간단한 분류 (간단한 작업 위주), semantic retrieval에 유능
  - Ada : 텍스트 구문분석, 단순 분류, 주소 수정
- Davinci는 품질이 좋은 대신 속도가 느림. 목적에 맞는 모델을 선택하는 것을 추천함.

LATEST MODEL	DESCRIPTION	MAX TOKENS	TRAINING DATA
`text-curie-001`	Very capable, faster and lower cost than Davinci.	2,049 tokens	Up to Oct 2019
`text-babbage-001`	Capable of straightforward tasks, very fast, and lower cost.	2,049 tokens	Up to Oct 2019
`text-ada-001`	Capable of very simple tasks, usually the fastest model in the GPT-3 series, and lowest cost.	2,049 tokens	Up to Oct 2019
`davinci`	Most capable GPT-3 model. Can do any task the other models can do, often with higher quality.	2,049 tokens	Up to Oct 2019
`curie`	Very capable, but faster and lower cost than Davinci.	2,049 tokens	Up to Oct 2019
`babbage`	Capable of straightforward tasks, very fast, and lower cost.	2,049 tokens	Up to Oct 2019
`ada`	Capable of very simple tasks, usually the fastest model in the GPT-3 series, and lowest cost.	2,049 tokens	Up to Oct 2019

https://platform.openai.com/docs/models/gpt-3

GPT-3 Family

✓ GPT-3.5

- GPT-3보다 긴 text를 받을 수 있음.
- GPT-3에 Code와 text를 이용하여 Fine-tuning함. (codex)
- GPT-3의 문제점
  - 안정성 고려 없이 Large-scale 데이터셋 기반으로 다음 단어를 예측하도록 수행함.
- 해결 : Fine-tuning
  - Instruction이용
  - Reinforcement Learning with Human Feedback (RLHF)를 도입함. (InstuctGPT)

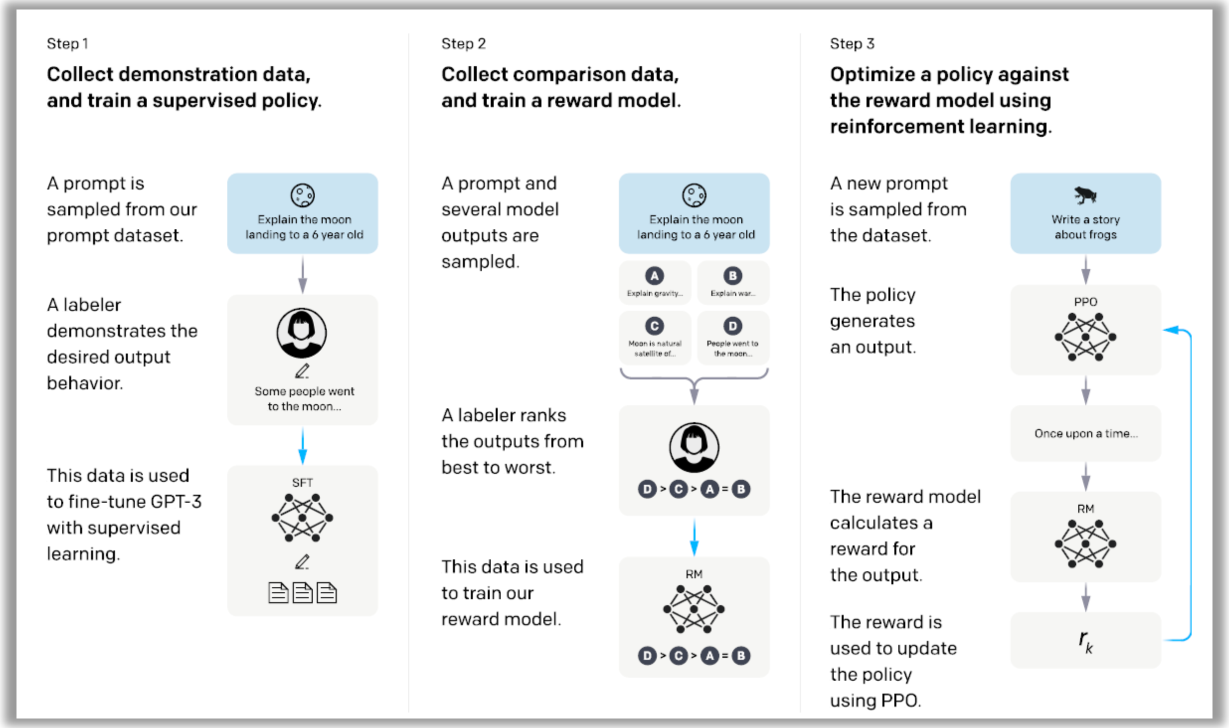
LATEST MODEL	DESCRIPTION	MAX TOKENS	TRAINING DATA
`gpt-3.5-turbo`	Most capable GPT-3.5 model and optimized for chat at 1/10th the cost of `text-davinci-003`. Will be updated with our latest model iteration.	4,096 tokens	Up to Sep 2021
`gpt-3.5-turbo-0301`	Snapshot of `gpt-3.5-turbo` from March 1st 2023. Unlike `gpt-3.5-turbo`, this model will not receive updates, and will be deprecated 3 months after a new version is released.	4,096 tokens	Up to Sep 2021
`text-davinci-003`	Can do any language task with better quality, longer output, and consistent instruction-following than the curie, babbage, or ada models. Also supports inserting completions within text.	4,097 tokens	Up to Jun 2021
`text-davinci-002`	Similar capabilities to `text-davinci-003` but trained with supervised fine-tuning instead of reinforcement learning	4,097 tokens	Up to Jun 2021
`code-davinci-002`	Optimized for code-completion tasks	8,001 tokens	Up to Jun 2021

https://platform.openai.com/docs/models/gpt-3

GPT-3 Family

✓ GPT-3.5 : Instruct-GPT

- Reinforcement Learning with Human Feedback
  1. Prompt를 보고 사람이 직접 답안(desired output)을 작성함.
     - (Prompt, desired output)을 학습
     - Supervised fine-tuning
  2. Prompt와 모델의 output들에 대해서 순위를 매김.
     - 사람이 어떤 output을 선호하는 지에 대해 reward model이 학습함.
  3. Prompt를 보고 추론한 결과(completion)를 reward model이 평가하여 reward를 계산함.
     - Reward를 InstructGPT에 주어지고, policy(instructGPT)를 업데이트하여 사람이 원하는 output에 가까운 결과를 냄.



https://openai.com/research/instruction-following

GPT-3 Family

✓ Instruct-GPT : Reinforcement Learning with Human Feedback

Prompt
8살 아이에게 이순신 장군에 대해 설명해보세요.

이순신 장군은 대한민국의 역사에서 매우 중요한 인물 중 하나입니다. 그는 조선 시대에 살았고, 조선의 왕과 함께 일본에 대항하는 전쟁을 치르며 유명해진 장군입니다.

Pre-trained GPT-3

Supervised fine-tuning

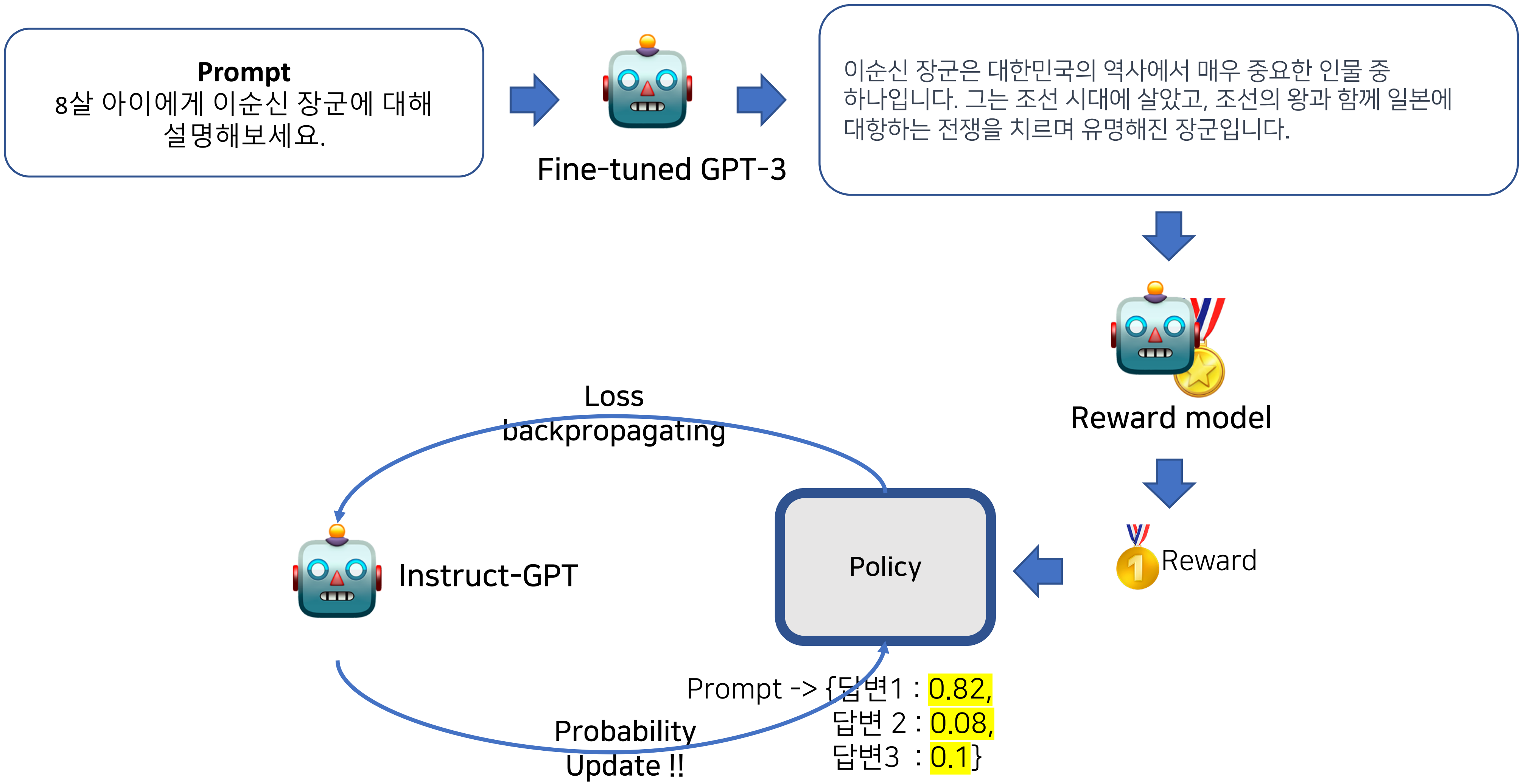
1. 이순신 장군은 15세기에 조선의 왕으로 즉위하여 1418년부터 1450년까지 32년 동안 다스리며 놀라운 업적을 이루었습니다.
2. 8살 어린아이에게는 이순신 장군의 업적을 이해하기에는 조금 어려울 수 있지만, 그의 용감한 모습과 국가를 위해 헌신한 모습을 간단하게 알려주면 좋을 것입니다.
3. 이순신 장군은 정말로 용감하고 뛰어난 전략가였습니다. 그는 조선의 왕과 함께 일본의 침략을 막기 위해 …
4. 이순신 장군은 우리나라의 역사에서 큰 영향을 미쳤으며 그의 인격과 지도력으로 많은 사람들에게 영감을 주었습니다.

3 > 4 > 2 > 1

Reward model

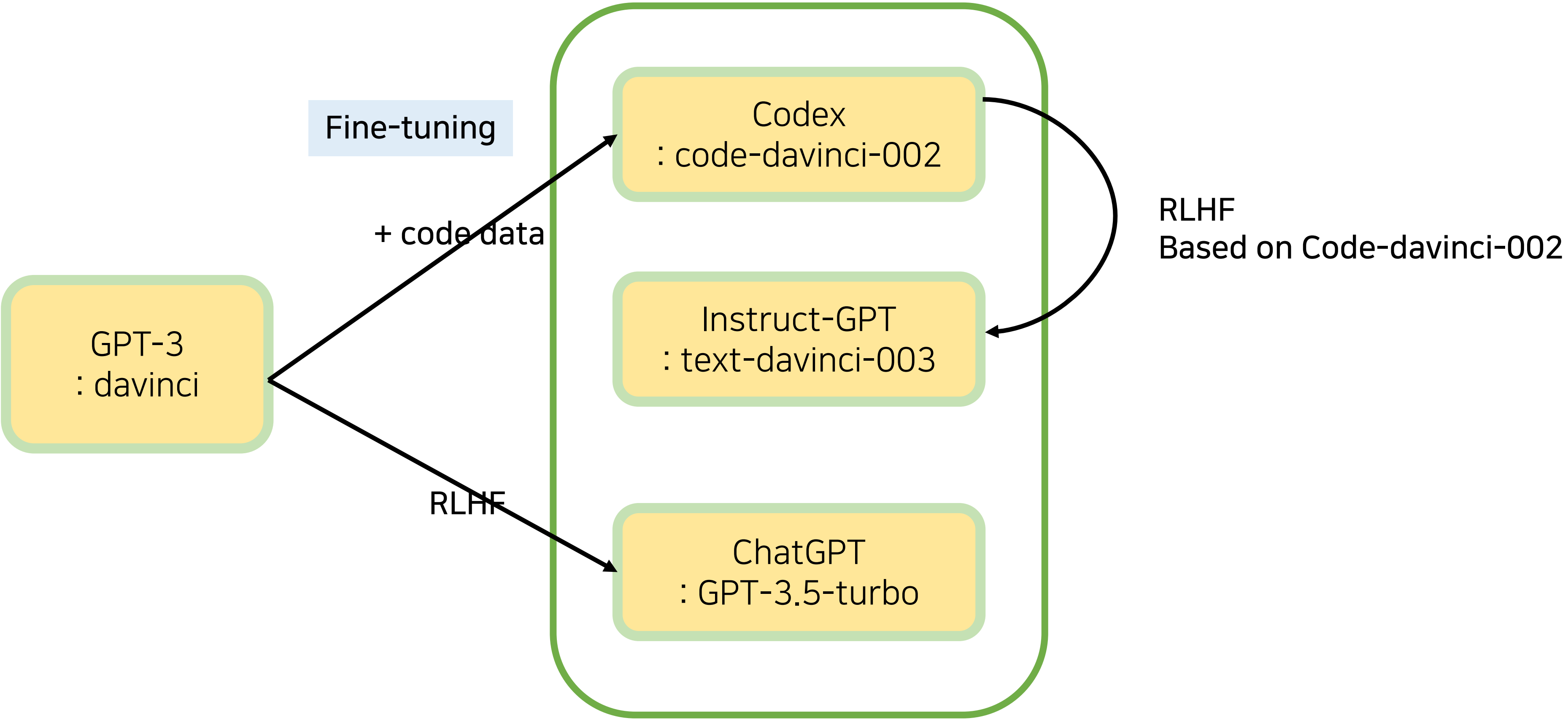
GPT-3 Family

✓ Instruct-GPT : Reinforcement Learning with Human Feedback



GPT-3 Family

✓ GPT-3, GPT-3.5, instruct-GPT와 ChatGPT의 관계

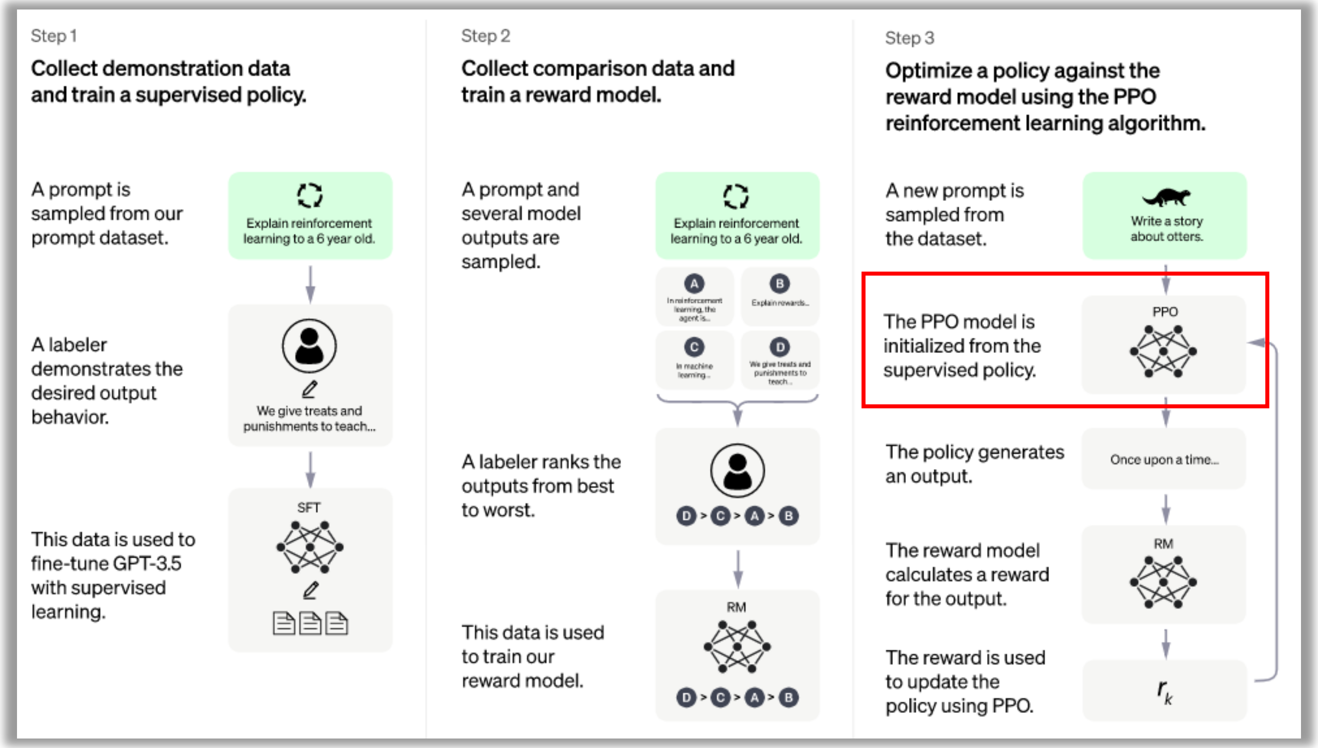


05 GPT-4

GPT-4

✓ Reinforcement Learning with Human Feedback

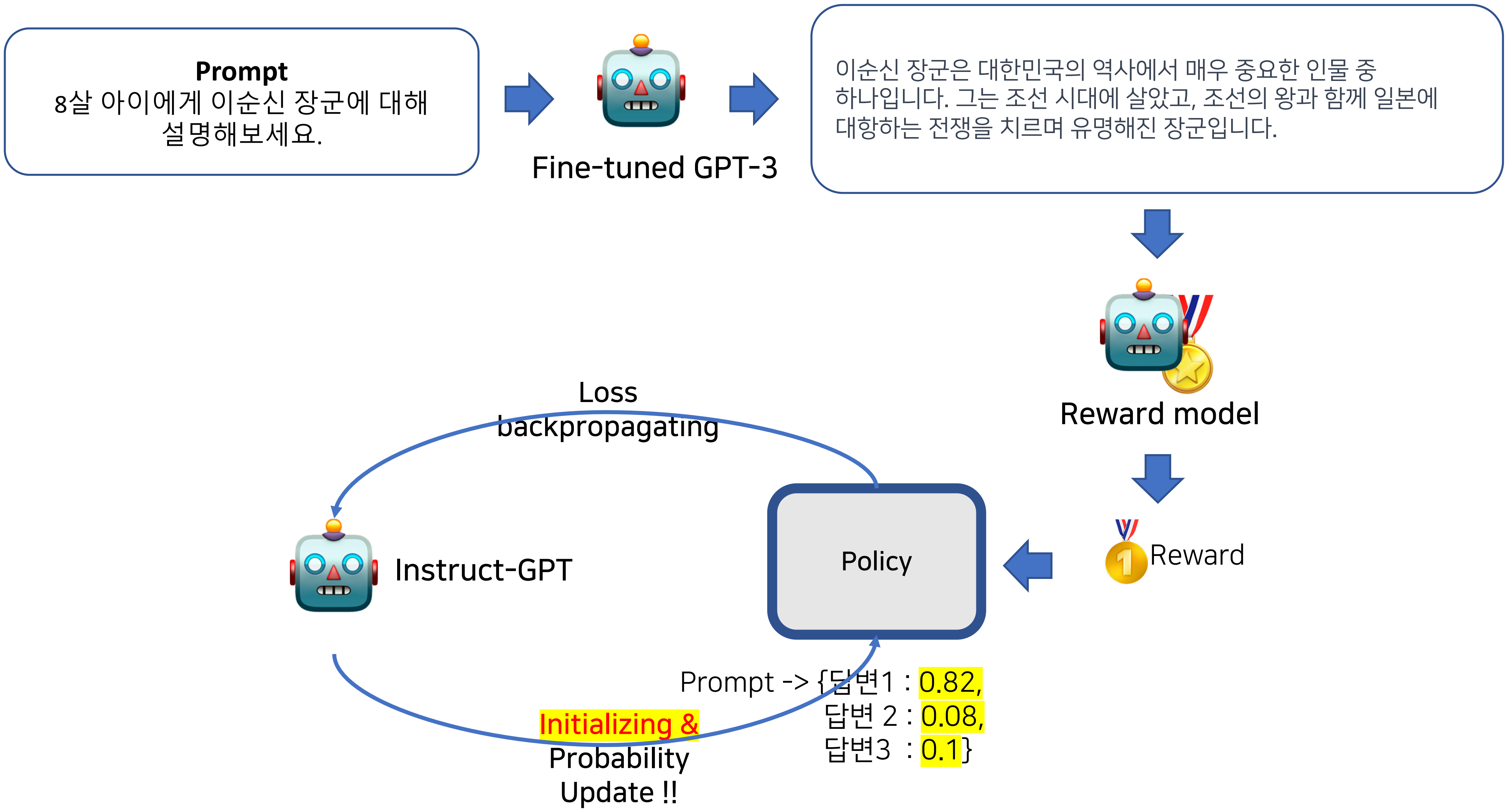
1. Prompt를 보고 사람이 직접 답안(desired output)을 작성함.
    - Prompt – desired output으로 지도학습
    - Supervised fine-tuning
2. Prompt와 모델의 output들에 대해서 순위를 매김.
    - 사람이 어떤 output을 선호하는 지에 대해 reward model이 학습함.
    - Reward 기준 : helpful, truthful, harmless
3. Prompt를 보고 추론한 결과(completion)를 reward model이 평가하여 reward를 계산함.
    - <u>Step1의 model로 부터 PPO 모델 생성함.</u>
    - PPO는 prompt기반으로 생성한 Output과 output에 따른 reward를 곱해 원하는 output에 대해 극대화 되도록 함.



https://openai.com/product/gpt-4

GPT-4

✓ GPT-4 : Reinforcement Learning with Human Feedback



GPT-4

✓ Capability

- Multi-modal : image, code, text 등
- GPT3.5 약 3,000단어 처리 가능 -> GPT4 25,000단어 처리 가능
- GPT-3.5보다 안정적이고 창의적이고 어려운 문제에 답변이 가능

Capabilities

In a casual conversation, the distinction between GPT-3.5 and GPT-4 can be subtle. The difference comes out when the complexity of the task reaches a sufficient threshold—GPT-4 is more reliable, creative, and able to handle much more nuanced instructions than GPT-3.5.

https://openai.com/research/gpt-4

GPT-4

✓ **Limitation**

✓ Hallucination

- 틀린 답변을 맞는 말처럼 생성하는 현상
- *we engaged over 50 experts from domains such as AI alignment risks, cybersecurity, biorisk, trust and safety, and international security to adversarially test the model.*
- 결과적으로 이전 보다 많이 줄었음

https://openai.com/research/gpt-4