

훈련 교 · 강사 보수교육

# 생성형 AI 이해

# 목차

## 1. 생성형 AI의 종류

- 생성 모델이란?

## 2. 생성형 언어 모델

- Word Embedding
- Transformer Model
- Chat-GPT

## 3. Prompt Engineering

- Prompt and Completion
- Zero shot Inference
- 상황 내 학습
- Good Prompt 작성 방법

## 4. 생성형 Image Model

- GAN
- Diffusion Model
- DALL-E
- Multi-modal model



# 생성형 AI의 종류



## 소목차

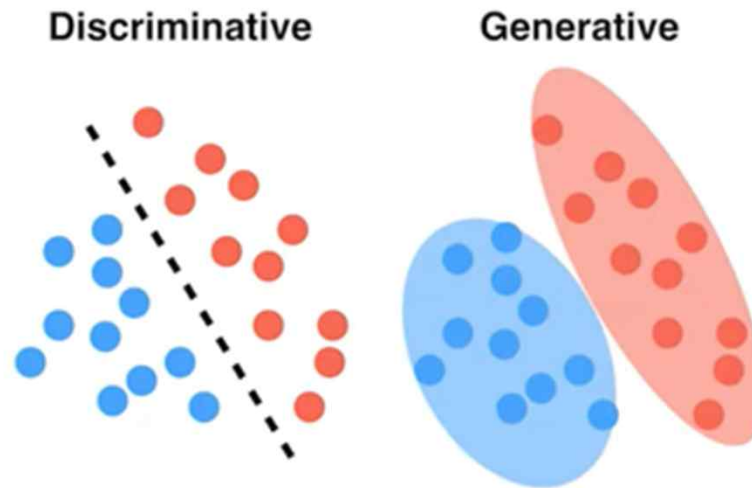
→ 생성 모델이란 ?

## 생성 모델이란 ?

- 실제 존재하지는 않지만 있을 법한 이미지를 생성할 수 있는 모델

- 분류/예측/군집화
- Feature와 label 간의 관계 학습

분류 모델



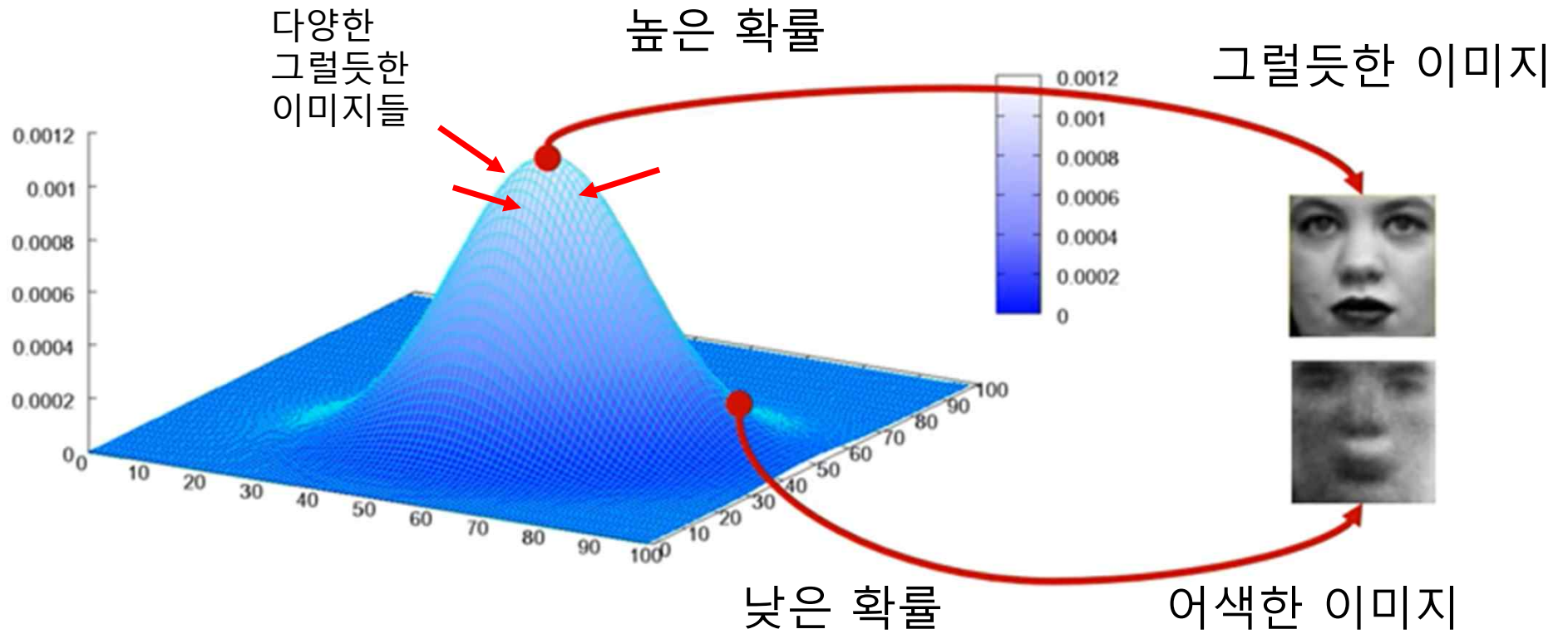
- 학습했던 데이터와 유사한 데이터 생성
- 데이터의 분포 학습
- 다음 단어 예측

생성 모델

결정 경계(Decision Boundary)를 학습

각 클래스의 분포(Distribution)를 학습

## 결합확률(Joint Probability)을 학습



## 생성형 AI의 종류

- Chatbot
- 가짜 Image 생성 (GAN)
- Text 로 부터 이미지 생성
- Code 생성
- Voice 생성
- Video 생성, etc

# 생성형 언어 모델

## Generative Language Model



## 소목차

- 최근의 NLP 발전 과정
- Word Embedding
- Transformer Architecture
- Large Language Model
- How to generate language ?
- Chat-GPT

## 최근의 NLP 발전 과정

- 2014 – RNN based sequence-to-sequence model 제안 (조현경 외)
- 2017 – **Transformer** (Google 연구진, "Attention is all you need") 논문 발표
- 2018. 6 – OpenAI GPT (Generative Pre-Training)
- 2018.11 – Google's BERT
- 2019. 2 – OpenAI GPT-2
- 2019.11 – Google's T5 (Text-To-Text Transfer Transformer)
- 2020. 5 – OpenAI GPT-3 → 1750 억 개의 parameter size
- 2022. 11 – ChatGPT (chatting version of GPT-3)
- 2023 – GPT-4, BARD, CLOVA-X
- 2024 - Gemini → **BIG, BIGGER, MORE BIGGER**



## Word Embedding (Feature 화 표시)

dimension	Man (5391)	Woman (9853)	King (4914)	Queen (7157)	Apple (456)	Orange (6257)
성별	-1	1	-0.95	0.97	0.00	0.01
귀족	0.01	0.02	0.93	0.95	-0.01	0.00
나이	0.03	0.02	0.7	0.69	0.03	-0.02
음식	0.04	0.01	0.02	0.01	0.95	0.97



**Man (5391) 의 4 dimension vector 표시**



**King (4914) 의 4 dimension vector 표시**

## Embedding matrix (example)

words

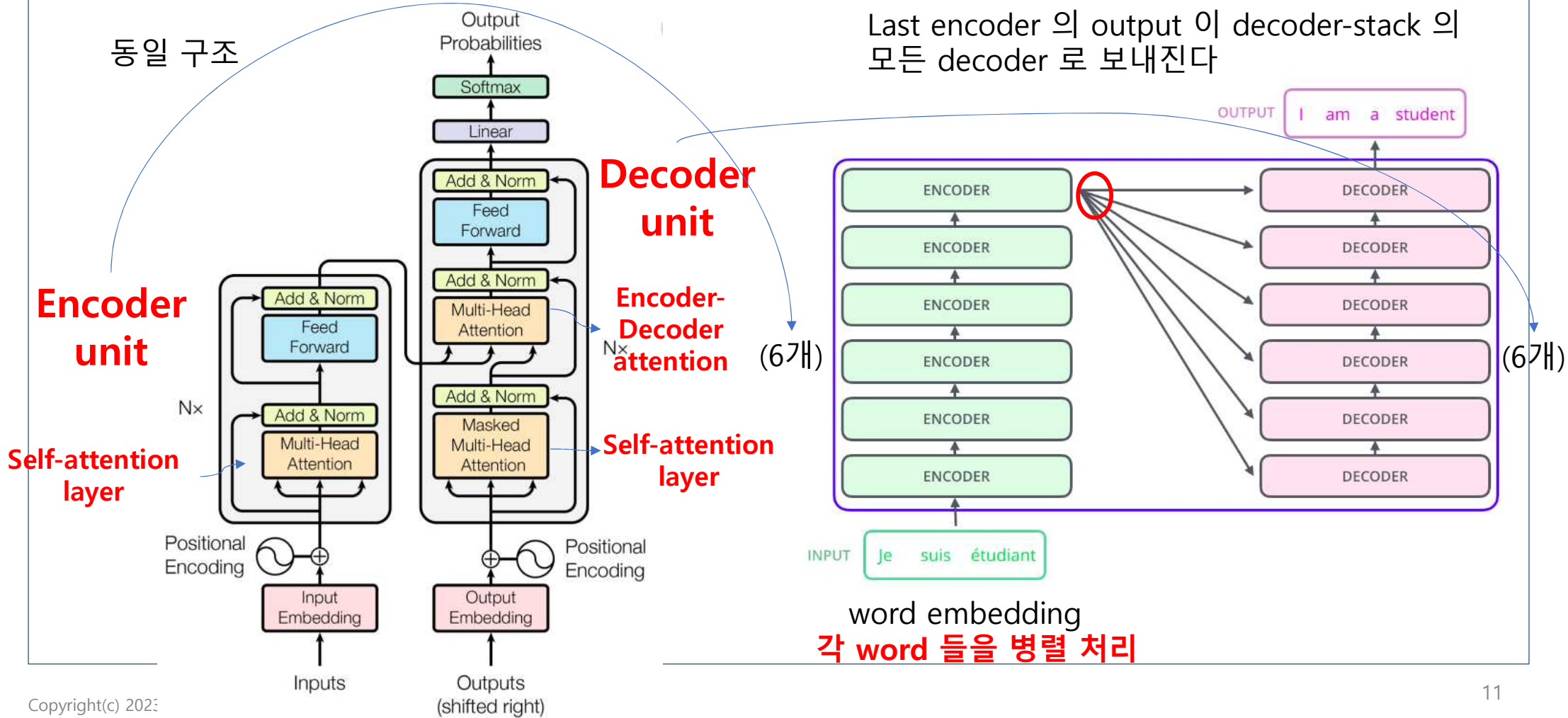
	0	1	2	3	4	5	6	7	8	9	...	290	291	292	
fox	-0.348680	-0.077720	0.177750	-0.094953	-0.452890	0.237790	0.209440	0.037886	0.035064	0.899010	...	-0.283050	0.270240	-0.654800	0.101
ham	-0.773320	-0.282540	0.580760	0.841481	-0.452890	0.237790	0.209440	0.037886	0.035064	0.899010	...	0.464470	0.481400	-0.829200	0.354
brown	-0.374120	-0.076264	0.109260	0.186621	-0.452890	0.237790	0.209440	0.037886	0.035064	0.899010	...	-0.015404	0.392890	-0.034826	-0.720
beautiful	0.171200	0.534390	-0.348540	-0.097234	0.101800	-0.170860	0.295650	-0.041816	-0.516550	2.117200	...	-0.285540	0.104670	0.126310	0.120
jumps	-0.334840	0.215990	-0.350440	-0.260020	0.411070	0.154010	-0.386110	0.206380	0.386700	1.460500	...	-0.107030	-0.279480	-0.186200	-0.540
eggs	-0.417810	-0.035192	-0.126150	-0.215930	-0.669740	0.513250	-0.797090	-0.068611	0.634660	1.256300	...	-0.232860	-0.139740	-0.681080	-0.370
beans	-0.423290	-0.264500	0.200870	0.082187	0.066944	1.027600	-0.989140	-0.259950	0.145960	0.766450	...	0.048760	0.351680	-0.786260	-0.360
sky	0.312550	-0.303080	0.019587	-0.354940	0.100180	-0.141530	-0.514270	0.886110	-0.530540	1.556600	...	-0.667050	0.279110	0.500970	-0.270
bacon	-0.430730	-0.016025	0.484620	0.101390	-0.299200	0.761820	-0.353130	-0.325290	0.156730	0.873210	...	0.304240	0.413440	-0.540730	-0.030
breakfast	0.073378	0.227670	0.208420	-0.456790	-0.078219	0.601960	-0.024494	-0.467980	0.054627	2.283700	...	0.647710	0.373820	0.019931	-0.030
toast	0.130740	-0.193730	0.253270	0.090102	-0.272580	-0.030571	0.096945	-0.115060	0.484000	0.848380	...	0.142080	0.481910	0.045167	0.050
today	-0.156570	0.594890	-0.031445	-0.077586	0.278630	-0.509210	-0.066350	-0.081890	-0.047986	2.803600	...	-0.326580	-0.413380	0.367910	-0.260
blue	0.129450	0.036518	0.032298	-0.060034	0.399840	-0.103020	-0.507880	0.076630	-0.422920	0.815730	...	-0.501280	0.169010	0.548250	-0.310
green	-0.072368	0.233200	0.137260	-0.156630	0.248440	0.349870	-0.241700	-0.091426	-0.530150	1.341300	...	-0.405170	0.243570	0.437300	-0.460
kings	0.259230	-0.854690	0.360010	-0.642000	0.568530	-0.321420	0.173250	0.133030	-0.089720	1.528600	...	-0.470090	0.063743	-0.545210	-0.190
dog	-0.057120	0.052685	0.003026	-0.048517	0.007043	0.041856	-0.024704	-0.039783	0.009614	0.308416	...	0.003257	-0.036864	-0.043878	0.000
sausages	-0.174290	-0.064869	-0.046976	0.287420	-0.128150	0.647630	0.056315	-0.240440	-0.025094	0.502220	...	0.302240	0.195470	-0.653980	-0.290
lazy	-0.353320	-0.299710	-0.176230	-0.321940	-0.385640	0.586110	0.411160	-0.418680	0.073093	1.486500	...	0.402310	-0.038554	-0.288670	-0.240
love	0.139490	0.534530	-0.252470	-0.125650	0.048748	0.152440	0.199060	-0.065970	0.128830	2.055900	...	-0.124380	0.178440	-0.099469	0.000
quick	-0.445630	0.191510	-0.249210	0.465900	0.161950	0.212780	-0.046480	0.021170	0.417660	1.686900	...	-0.329460	0.421860	-0.039543	0.150

학습된 300 차원 featrues

20 rows x 300 columns

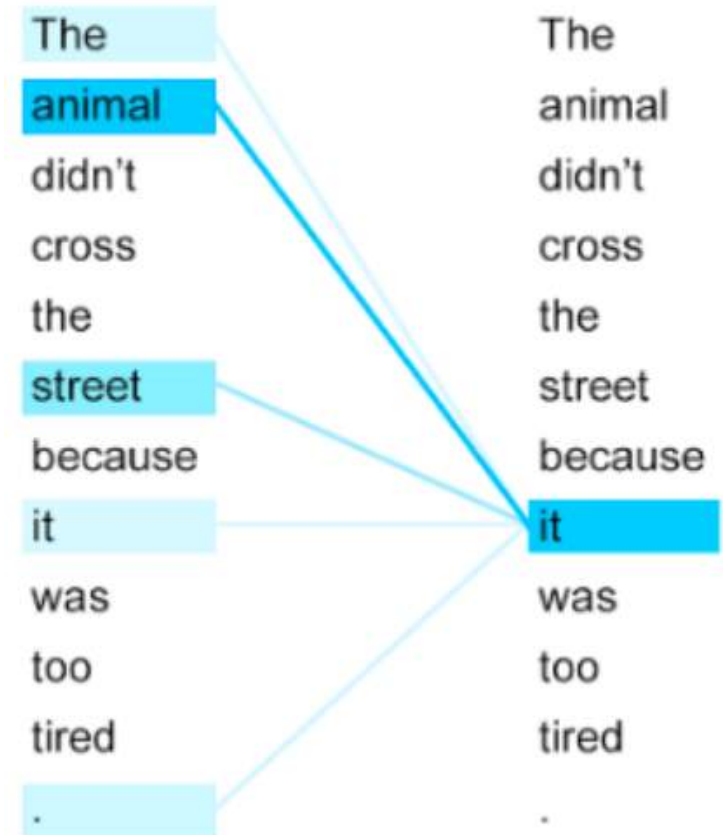
학습된 300 차원 features

# Transformer Architecture

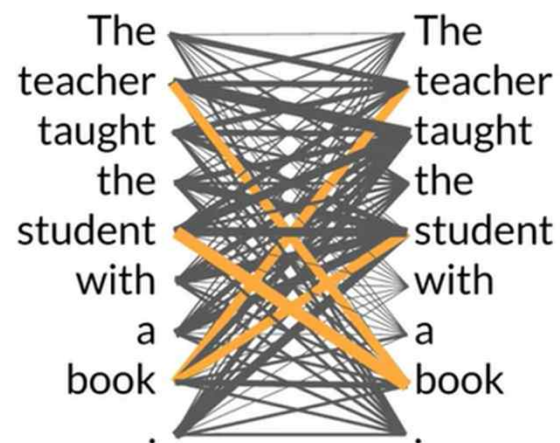
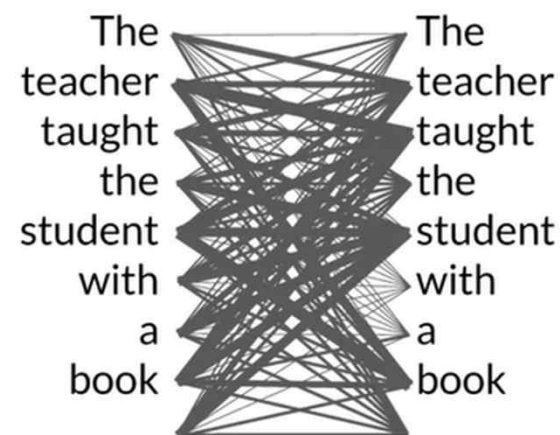
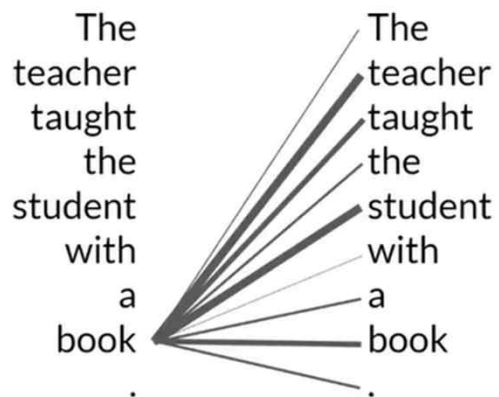
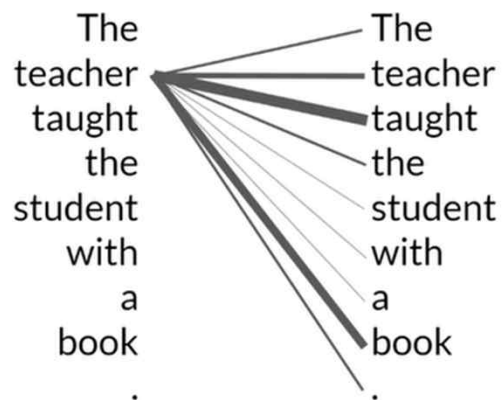


## Self-Attention (Intra-Attention)

- Attention 을 자기 자신에 대해 수행  
→ 문장 내 단어들 간의 유사도를 구함
- Self-attention 계산  
3 개의 vector 필요 (훈련 과정에서 스스로 학습)
  - Query Vector
  - Key Vector
  - Value Vector

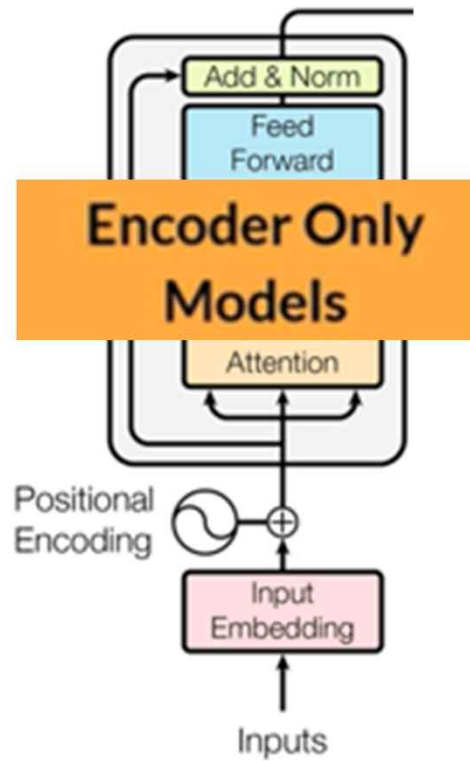


## Self-Attention (Intra-Attention)

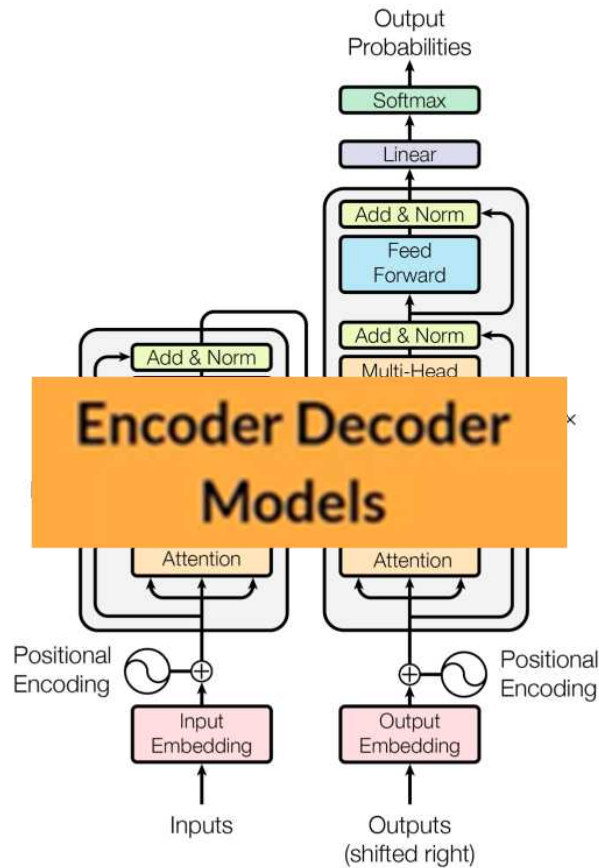




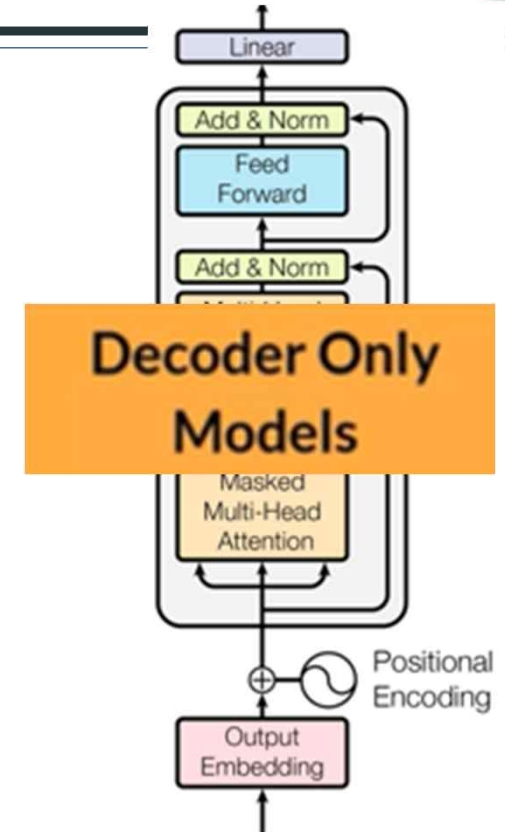
## Transformers 파생 모델



- BERT



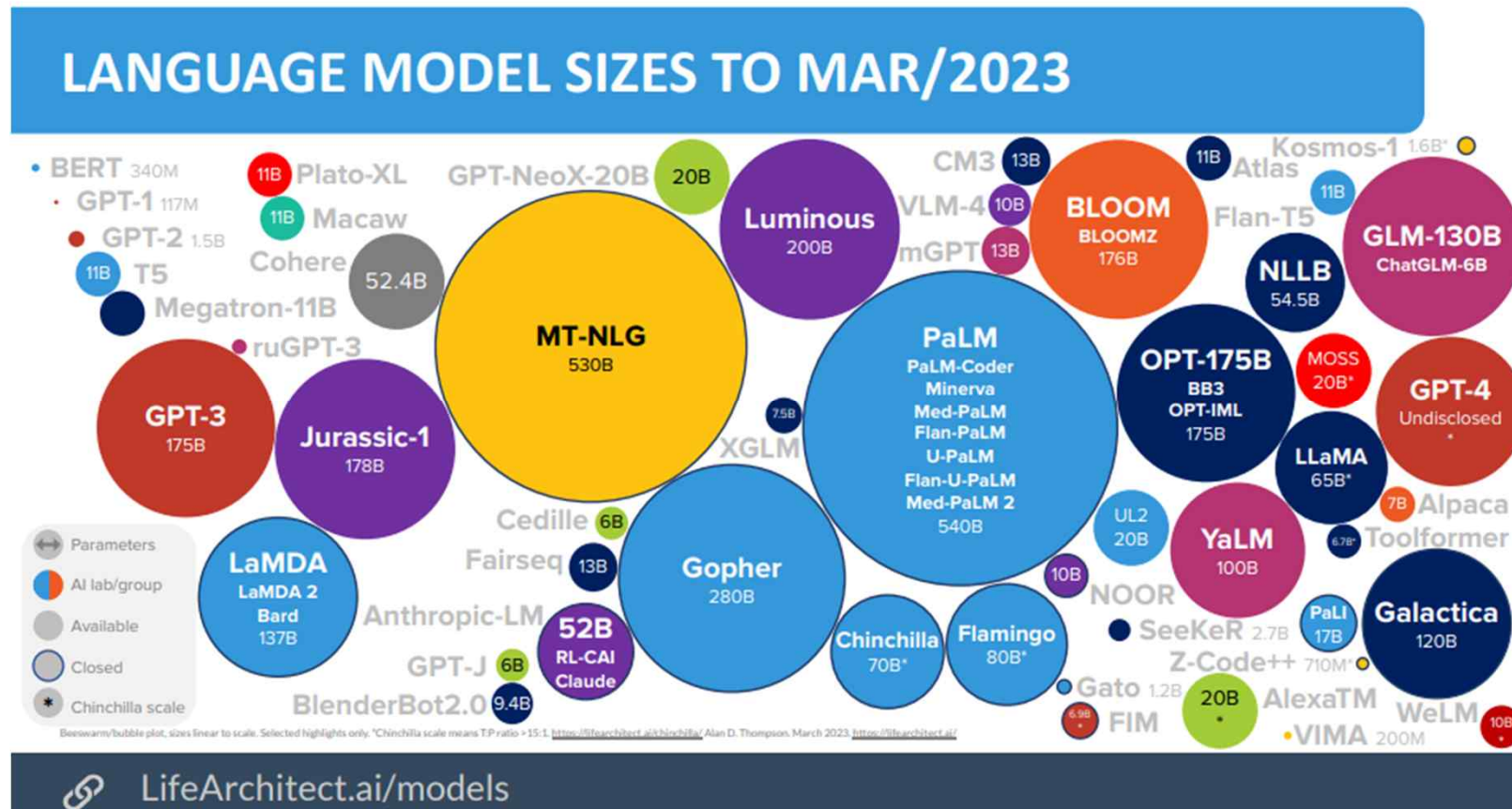
- T5
- BART



- GPT 모델 제품군
- BLOOM
- Jurassic
- LLaMA

## Large Language Model (LLM)

- language model – next word prediction model



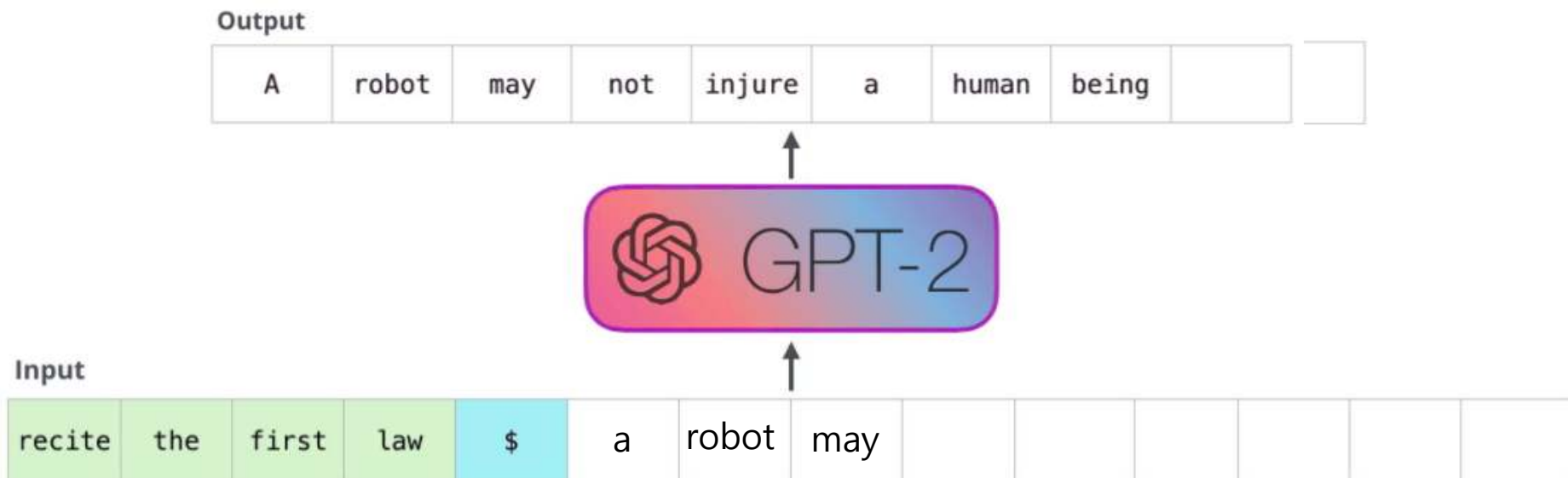
## LLM 응용 분야

- Chatbot - 대화
- 글쓰기 - Text generation (문장 생성)
- 요약 - 긴 문장 summary
- 번역 - Machine Translation
- 프로그램 코딩
- 개체명 인식(Entity Recognition)
- Sentiment Analysis (감성분석) - ex) 이 영화평이 긍정적인가 부정적인가 ?
- Topic modeling (주제분류) - ex) news 내용을 정치, 스포츠, 과학등으로 분류
- 형태소 분석
- 철자법 교정 (Spelling correction)
- Spam Detection : email 이 spam 인지 ham 인지 구분



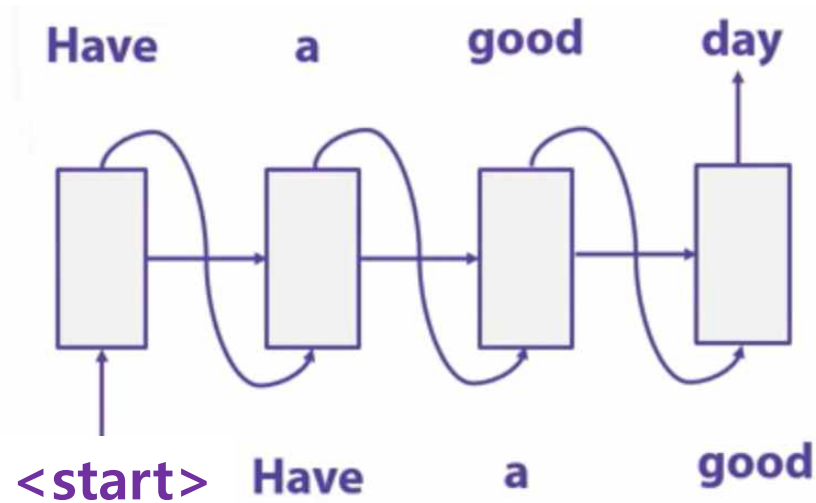
## GPT 는 auto-regressive 한 입력 구조

- 각 토큰이 생성 된 후 해당 토큰이 입력 시퀀스에 추가
- 그 새로운 시퀀스는 다음 단계에서 모델에 대한 입력이 된다



## How to generate language ?

- 이전 time step 의 output 을 next time step 의 input 으로 feed 하고, 각 step 에서 가장 높은 확률의 다음 단어를 선택 (greedy selection)하거나 혹은 확률 분포에 따라 sampling

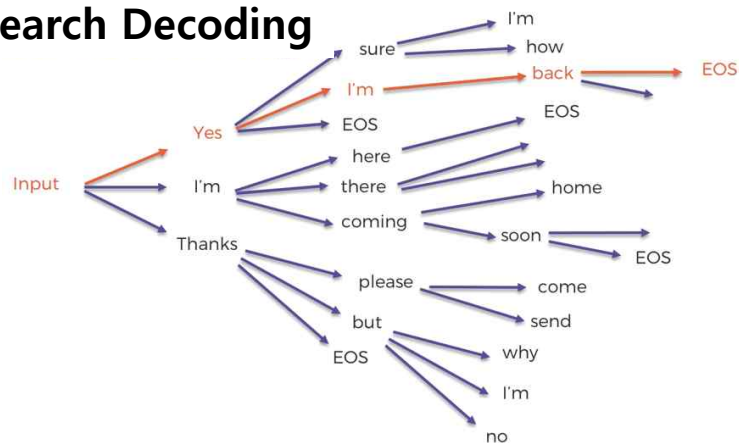


## Decoding Strategy

- Greedy 전략
  - softmax 분포 중 가장 높은 확률 (argmax) 선택, 언제나 같은 번역
  - $\text{argmax } P(\text{Over the line !} \mid \text{Çizgiyi geçtin !})$
- Sampling 전략 (사후확률분포)
  - 분포 확률에 따라 Random sampling, 매번 번역이 바뀔 수 있음
  - `np.random.choice(len(probs), p=probs)`
- Beam-search 전략
  - 단순히 첫번째 단어를  $\text{argmax}$  로 선택하면 1스텝 에서라도 문법 상 실수를 할 경우, 전체 문장의 번역에 큰 실수가 되므로, 각각의 타임스텝  $t$  마다  $b$  개의 sequence 후보 군을 유지

## Highest joint probability

### Beam Search Decoding



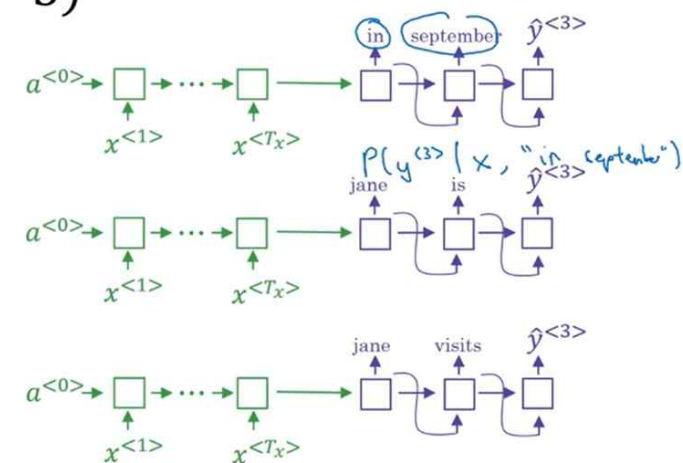
### Beam search ( $B = 3$ )

in september

jane is

jane visits

$$P(y^{<1>, y^{<2>} | x)$$



jane visits africa in september. <EOS>

Andrew Ng

## What is ChatGPT ?

- ChatGPT라고 불리는 GPT-3.5는 GPT-3의 작지만 대화에 특화된 버전.  
이 모델은 채팅의 맥락, 구조 및 어조를 더 깊이 이해하면서 자연스럽게 대화를 이어가는 유일한 목적으로 훈련. 따라서 채팅과 관련하여 범용 GPT-3에 비해 훨씬 더 정확하고 빠른 모델을 만든다.
- GPT-4는 ChatGPT의 진화된 Multi-Modal 버전.  
텍스트나 이미지 입력을 받아 텍스트를 출력하는 대규모 다중 모드 모델로, 이전 모델들보다 더 넓은 일반 지식과 진보된 추론 능력을 통해 더 정확하게 어려운 문제를 해결할 수 있다.

## ChatGPT 사용 example



나는 인공지능 강사야. 나의 인지도를 높이기 위한 방법 5 가지만 알려줘. 한글로 알려줘

### Zero shot 추론



1. 최신 인공지능 개발을 추적하기: 산업 뉴스와 발전, 회의와 워크숍, 관련 연구 논문과 블로그를 읽어보세요.
2. 인공지능의 특정 분야에 전문화하기: 컴퓨터 비전, 자연어 처리, 보상 학습 등의 특정 인공지능 분야에 집중하고 그 분야에서 전문가가 되세요.
3. 자신만의 인공지능 프로젝트 구축하기: 실제 경험은 인공지능의 이해를 깊이 하는 최고의 방법 중 하나입니다. 간단한 프로젝트부터 시작하고 경험이 많아질수록 복잡성을 증가시킵니다.
4. 다른 인공지능 전문가와 협업하기: 온라인 커뮤니티와 포럼에 가입하고, 미팅에 참석하고, 다른 인공지능 전문가와 프로페셔널과 협업할 기회를 찾아보세요.
5. 다른 사람들에게 인공지능

## ChatGPT 사용 example



소프트웨어 산업협회에서 올해 처음으로 인공지능 트렌드 과정을 열었고 많은 사람들의 호평 속에 끝났습니다. **이걸 홍보하려는데 기사로 써 주세요.**

### 기사 작성 요청



"인공지능의 미래를 예측하고 현재의 최신 기술을 이해하는 것이 중요한데, 소프트웨어 산업협회에서 이에 대한 첫 번째 기회를 제공하였습니다. 올해 처음으로 열린 인공지능 트렌드 과정은 많은 사람들의 호평 속에 끝나었습니다. 트렌드 과정에서는 최신 인공지능 기술과 이를 활용한 사례, 인공지능의 미래 전망 등에 대해 심도 있게 논의하였습니다. 이 과정은 인공지능을 활용한 미래 기술의 발전에 대한 이해를 높이는 데 도움이 될 것입니다."

## ChatGPT 사용 example



100 이하의 숫수를 반환하는 python code 를 작성해줘

### Program Code 작성



100 이하의 소수를 반환하는 Python 코드를 작성해 드리겠습니다. 소수는 1과 자기 자신만으로 나누어 떨어지는 양의 정수를 말합니다. 📄 👍 👎

python

Copy code

```
def is_prime(n):  
    if n <= 1:  
        return False  
    elif n <= 3:  
        return True  
    elif n % 2 == 0 or n % 3 == 0:  
        return False  
    i = 5  
    while i * i <= n:  
        if n % i == 0 or n % (i + 2) == 0:  
            return False  
        i += 6  
    return True
```

```
def find_primes(limit):  
    primes = []  
    for num in range(2, limit + 1):  
        if is_prime(num):  
            primes.append(num)  
    return primes  
  
limit = 100  
prime_numbers = find_primes(limit)  
print(prime_numbers)
```



# Prompt Engineering (Prompt Optimization)



## 소목차

---

- Prompt 관련 용어
- Prompting and Completion
- zero shot prompting(inference)
- 상황 내 학습(in-context learning)
- COT(Chain-of-thought) Prompting
- Good Prompt 작성 방법

## Prompting and Completion

### Prompt

Prompt Engineering이  
뭔가요?



Model



### Completion

Prompt engineering은 인공지능,  
특히 자연어 처리를 하는 AI 모  
델에게 효과적인 질문이나 지  
시를 제공하는 기술과 방법을  
의미합니다.

## Prompt 관련 용어

- Prompt - 모델에게 주는 가이드 (입력 텍스트)
  - Instruction (명령) - prompt의 핵심. 텍스트 요약, 번역, 분류 등
  - Context - instruction을 구체화 할 수 있는 추가 정보.
  - Input Data - AI 모델이 지시한 작업을 완료하기 위해 처리할 정보
  - Output Indicator - 암시적 or 명시적
- 추론 - 텍스트를 생성하는 행위
- Completion - 출력 텍스트
- Context window - 프롬프트에 사용할 수 있는 전체 텍스트 양 또는 메모리
- 프롬프트 엔지니어링 - 프롬프트를 개발하고 개선하는 작업
  - 나은 결과를 생성하도록 하는 모델이 수행하기를 원하는 작업의 예를 프롬프트 내에 포함
  - 상황 별 예제를 제공

## Prompt Engineering 예

- **Instruction (명령):**  
"다음 내용을 요약해 주세요."
- **Context (문맥):**  
"다음 내용은 민주주의와 공산주의에 대한 비교입니다. 중요한 사항들을 명확하게 이해할 수 있도록 주요 내용을 간략히 정리해야 합니다."
- **Input Data:**  
"공산주의는 생산수단의 .....(중략) ..... 제2차 세계대전의 경로를 통하여 동유럽 및 아시아의 공산주의권이 성립하기 시작했다. 그러나, 1991년 소비에트 연방이 붕괴한 이후부터 공산주의 운동은 사실상 쇠퇴의 일로를 맞고 있다"
- **Output Indicator:**  
"민주주의와 공산주의 비교의 핵심 포인트를 표로 정리해 주세요."



You

다음 내용을 요약해 주세요.

**Instruction (명령)**

다음 내용은 민주주의와 공산주의에 대한 비교입니다. 중요한 사항들을 명확하게 이해할 수 있도록 주요 내용을 간략히 정리해야 합니다.

**Context (문맥)**

공산주의(라틴어: Communismus, 독일어: Kommunismus, 러시아어: Коммунизм, 영어: Communism, 共產主義)는 생산수단의 공공 소유에 기반을 둔 경제·사회·정치 공동체 형성에 관한 사상 또는 이러한 사회 형성을 목표로 삼는 형식적·실질적 정치 운동에 관한 사상이다.

플라톤(Platon)의 『국가』를 통해서 알 수 있듯이, 본래 공산주의적 사고는 전근대 사회에

(중략)

했다. 그러나, 1991년 소비에트 연방이 붕괴한 이후부터 공산주의 운동은 사실상 쇠퇴의 일로를 맞고 있다.

**Input Data**

민주주의와 공산주의 비교의 핵심 포인트를 표로 정리해 주세요.

**Output Indicator**



ChatGPT

민주주의와 공산주의의 핵심 비교 포인트는 다음과 같습니다:

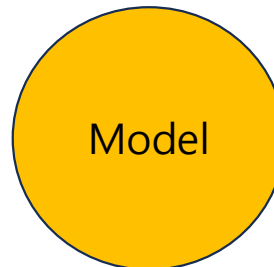
특징	민주주의	공산주의
기원 및 사상	국민에 의한 정치 권력 선출, 자유와 평등을 기반으로 하는 다양한 정치적 사상	플라톤과 마르크스의 이론에 영향을 받은 생산수단의 공공 소유 및 사회 평등 추구
경제 체제	주로 시장 경제 및 자본주의적 원칙에 기반	중앙 계획 경제, 생산수단의 공동체 또는 국가 소유

## zero shot prompting(inference)

- 답을 나타내는 예시나 입력 데이터 없이 답변 생성
- 초보자의 사용 방식 (단순 질문 방식) – 모델의 기존 지식에 의존
- Guide가 없으므로 원하는 답이 아닐 가능성 높음

### Prompt

미래의 자동차가 어떤 모습  
일지 상상해 보세요?



### Completion

미래의 자동차는 현재의 기술  
과 추세를 기반으로 상당히  
진보된 모습을 가질 것입니다.  
몇 가지 가능성은 다음과 같  
습니다:  
(이하 생략)



You

미래의 자동차가 어떤 모습일지 상상해 보세요

## Zero-shot Prompting (Inference)



ChatGPT

미래의 자동차는 현재의 기술과 추세를 기반으로 상당히 진보된 모습을 가질 것입니다. 몇 가지 가능성은 다음과 같습니다:

1. **완전 자율 주행:** 미래의 자동차는 고도로 발전된 자율 주행 기능을 탑재하여 운전자의 개입 없이 도로에서 안전하게 운행할 수 있습니다. 이러한 자동차는 교통 사고를 크게 줄이고, 도로 효율성을 높이며, 운전자에게 여가 시간을 제공합니다.
2. **친환경 기술:** 전기차와 수소 연료 전지차 같은 친환경 기술은 더욱 보편화되어 환경 오염을 크게 줄일 것입니다. 이는 지속 가능한 교통의 미래를 위한 중요한 단계입니다.
3. **통합된 연결성:** 차량 내부 기술은 스마트 홈, 스마트폰, 사무실 등과 완벽하게 통합됩니다. 운전자와 승객은 이동 중에도 연결성을 유지하고, 다양한 엔터테인먼트와 업무 관련 기능을 이용할 수 있습니다.
4. **개인 맞춤형 디자인:** 미래의 자동차는 더욱 개인화되어 사용자의 취향과 필요에 맞게 맞춤 설정이 가능해질 것입니다. 외부 디자인부터 내부 기능, 심지어 주행 스타일까지 사용자의 선호에 따라 조정될 수 있습니다.

## 상황 내 학습(in-context learning) – one shot inference

- 모델에게 특정 작업이나 개념에 대해, 자신이 원하는 한가지 예(example, shot)를 instruction 과 함께 제시
- 모델은 제공된 예와 유사한 새로운 데이터를 생성 혹은 분류

### Prompt



영화평을 다음과 같이 긍정적 또는 부정적으로 감성 분석하세요.

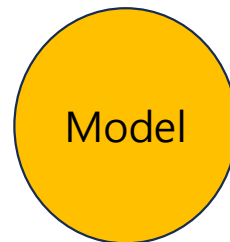
영화평: 이 영화 정말 짱이야!

감성: 긍정적

영화평: 보다가 지루해서 죽을 뻔 했어!

감성:

One  
example



### Completion



영화평: 보다가 지루해서 죽을 뻔 했어!

감성: 부정적





영화평을 다음과 같이 긍정적 또는 부정적으로 감성 분석하세요.

영화평: 이 영화 정말 짱이야!

감성: 긍정적

영화평: 보다가 지루해서 죽을 뻔 했어!

감성:

In-Context Learning (One shot Inference)  
One example



영화평: 보다가 지루해서 죽을 뻔 했어!

감성: 부정적



This movie was great!"는 긍정적인 리뷰입니다.

그렇다면 "I really enjoyed this film!"은 어떤 리뷰인가요?"

In-Context Learning  
One example



"I really enjoyed this film!"는 긍정적인 리뷰입니다. "really enjoyed"라는 표현은 영화에 대한 긍정적인 반응을 나타내기 때문입니다.

## 상황 내 학습(in-context learning) – few shot inference (prompting)

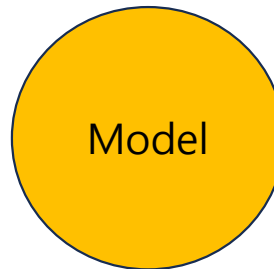
- 모델에게 특정 작업이나 개념에 대해, 자신이 원하는 몇가지 예(example, shot)를 instruction 과 함께 제시
- 모델은 제공된 예와 유사한 새로운 데이터를 생성 혹은 분류

### Prompt

다음은 리뷰를 분류하세요.  
나는 이 영화가 좋아요.  
감정: 긍정적

다음은 리뷰를 분류하세요.  
나는 이 의자가 싫어요.  
감정: 부정적

다음은 리뷰를 분류하세요.  
나는 이 음악이 싫어요.  
감정:



### Completion

다음은 리뷰를 분류하세요.  
나는 이 영화가 좋아요.  
감정: 긍정적

다음은 리뷰를 분류하세요.  
나는 이 의자가 싫어요.  
감정: 부정적

다음은 리뷰를 분류하세요.  
나는 이 음악이 싫어요.  
감정: 부정적



'This movie was great!'와 'I loved the play'는 긍정적인 리뷰입니다.  
반면, 'I did not like the film'와 'The book was boring'는 부정적인 리뷰입니다.  
그렇다면 'The concert was amazing!'은 어떤 리뷰인가요?"

In-Context Learning  
Few examples



'The concert was amazing!'는 긍정적인 리뷰입니다. "amazing"이라는 단어는 강력한 긍정적  
인 감정을 표현하므로 이 리뷰는 콘서트에 대한 매우 긍정적인 반응을 나타냅니다.



질문: 스페인의 수도는 어디인가요?  
답: 마드리드  
질문: 이태리의 수도는 어디인가요?  
답: 로마  
질문: 프랑스의 수도는 어디인가요?  
답:

In-Context Learning  
Few examples



파리

## COT(Chain-of-thought) Prompting

- LLM의 추론 능력 개선을 위해 Google 연구원들이 고안
- 문제를 단계별로 분해하여 복잡한 추론 문제 해결  
(<https://arxiv.org/pdf/2201.11903v1.pdf>)
- Chain-of-thought prompt engineering은 다음과 같은 단계로 수행됩니다.
  1. 문제를 여러 단계로 나눕니다.
  2. 각 단계에 대해 LLM에게 질문하거나 수행할 작업을 지시하는 프롬프트를 작성합니다.
  3. LLM이 출력한 결과를 다음 단계의 프롬프트에 입력합니다.
  4. 이 과정을 반복하여 문제를 해결합니다.

## COT(Chain-of-thought) Prompting



나는 시장에 가서 10개의 사과를 샀다. 이웃에게 사과 2개를 주었고 수리공에게 2개를 주었다.  
나는 다시 가서 5개를 더 사서 1개를 먹었다. 그러면 지금 나는 몇개의 사과를 가지고 있는가?  
단계별로 생각해.



1. 당신은 처음에 시장에서 10개의 사과를 샀습니다.
  - 현재 사과 갯수: 10개
2. 그 후 이웃에게 2개의 사과를 주었습니다.
  - 현재 사과 갯수:  $10\text{개} - 2\text{개} = 8\text{개}$
3. 수리공에게도 2개를 주었습니다.
  - 현재 사과 갯수:  $8\text{개} - 2\text{개} = 6\text{개}$
4. 당신은 다시 시장에 가서 5개의 사과를 더 샀습니다.
  - 현재 사과 갯수:  $6\text{개} + 5\text{개} = 11\text{개}$
5. 마지막으로 1개의 사과를 먹었습니다.
  - 현재 사과 갯수:  $11\text{개} - 1\text{개} = 10\text{개}$

따라서, 당신은 현재 10개의 사과를 가지고 있습니다.

단계별 답변 요구

- 모델의 추론 과정을 파악할 수 있음.
- 사전 지식 없이 모델에 창의적인 자유를 부여

## 최신 생성형 LLM의 COT 예



영희는 10마리의 개를 키운다. 개 한마리를 산책시키고 돌보는데 0.5 시간이 필요하다. 갑자기 여행가는 친구가 개 두마리를 맡기고 떠났다. 그런데 이 개들 중 2마리가 각각 한마리씩 새끼를 낳았다. 새끼 한마리는 돌보는데 1시간이 필요하다. 이제 영희는 모든 개들을 돌보는데 일주일에 몇시간이 필요한가?

### GPT-4: 정답



영희가 키우던 개는 10마리이고, 친구의 개 2마리를 추가로 맡아 총 12마리의 개를 돌보고 있습니다. 이 개들 중 2마리가 각각 새끼를 한 마리씩 낳아, 추가로 새끼 개 2마리가 더해져 총 14마리의 개를 돌보아야 합니다.

1. 영희가 원래 키우던 개와 친구의 개를 돌보는데 필요한 시간:  $12\text{마리} \times 0.5\text{시간/마리} = 6\text{시간}$
2. 새끼 개를 돌보는데 필요한 시간:  $2\text{마리} \times 1\text{시간/마리} = 2\text{시간}$

따라서 영희는 하루에 모든 개를 돌보는데 총 8시간이 필요합니다.

일주일에는 필요한 시간은:  $8\text{시간/일} \times 7\text{일/주} = 56\text{시간/주}$  입니다.

COT형식의  
답변 출력

## Prompt Engineering(Optimization) 요약 - Good Prompt 작성 방법

- 좋은 prompt 를 만드는 세가지 전략
  - Zero-shot prompt : 예시 없는 prompt
  - Few-shot prompt : 몇 개의 예시를 보여 주며 질문
  - COT (Chain-of-thought) : Step-by-step 설명 요구
- Clear, non ambiguous task – LLM이 달성해야 할 구체적인 목표, 목적 설정
- 같은 prompt로 여러 결과를 만들고 좋은 결과물을 선택  
→ 같은 질문 여러 번 할 때의 시간/비용 trade-off





고위 DevOps 엔지니어 직책을 채용하기 위한 기술 인터뷰 질문 목록을 생성해줘

일반적인 prompting



고위 DevOps 엔지니어 직책을 채용하기 위한 인터뷰에서는, 캔디데이트가 복잡한 인프라와 시스템을 관리하고, 팀을 이끄는 능력, 그리고 문제를 해결하는 능력을 보유하고 있는지를 확인하는 것이 중요합니다. 다음은 적절한 기술 인터뷰 질문들입니다:

1. CI/CD 파이프라인을 설명하고, 어떻게 설계하고 구현하는지 예를 들어 설명해주세요.
2. 소스 코드에서 운영 환경까지의 워크플로우를 어떻게 자동화하나요?



조직 문화가 빠르게 변화하고 클라우드 기술로 개발 중인 기술 스타트업의 고위 DevOps 엔지니어 직책을 채용하기 위한 기술 인터뷰 질문 목록을 생성해줘

Better prompting  
- 구체적인 목표,  
목적 설정



DevOps 엔지니어 직책을 채용하기 위한 기술 인터뷰 질문은 다음과 같습니다:

1. 클라우드 기반 서비스를 활용한 이전의 프로젝트를 소개해주세요. 그 프로젝트에서 당신의 역할은 무엇이었나요?
2. AWS, Google Cloud, Azure와 같은 클라우드 서비스를 사용한 경험이 어떠한가요? 그 중 어떤 서비스를 가장 선호하나요?
3. 새로운 서버를 빠르게 배포하고, 효과적으로 모니터링하고, 이상을 발견하면 신속하게 대응하는 과정을 설명해 주시겠어요?





우리 e-commerce website 의 사용자 경험을 개선해줘

일반적인 prompting



사용자 경험(UX)을 개선하는 것은 웹사이트의 전반적인 성능을 향상시키는 데 중요한 요소입니다. 다음은 e-commerce 웹사이트의 사용자 경험을 개선하기 위한 몇 가지 제안입니다:

1. **속도 최적화:** 웹사이트의 로딩 속도는 사용자 경험에 매우 중요한 역할을 합니다. 사용자들은 웹사이트가 빠르게 로딩되기를 원하며, 그렇지 않으면 다른 사이트로 이동할 가능성이 높습니다. 따라서 이미지 최적화, 코드 최적화, CDN 사용 등을 통해 웹사이트의 속도를 최적화하시는 것을 추천드립니다.
2. **사용자 인터페이스(UI) 개선:** 사용자가 웹사이트를 쉽게 탐색할 수 있도록 직관적



고객 만족도 및 판매 전환율을 높이기 위해 우리 e-commerce website 의 사용자 경험을 개선할 수 있는 특별한 방법을 제안해 줘

Better prompting  
- 구체적인 목표,  
목적 설정



기본적인 사용자 경험 개선 방안 외에도, 전자 상거래 웹사이트를 위해 특별하게 고려해 볼 수 있는 몇 가지 전략들이 있습니다:

1. **증강현실(AR) 기능 도입:** 증강현실 기능을 사용하여 사용자들이 제품을 실제로 보고 느낄 수 있게 해주면 구매 결정을 돕습니다. 예를 들어, 가구를 판매하는 회사는 사용자가 자신의 집에서 어떻게 보일지 보여주기 위해 AR 기능을 사용할 수 있습니다.
2. **AI 챗봇 활용:** AI 챗봇은 고객 서비스를 개선하고, 고객이 문제를 빠르게 해결할 수 있도록 도와줄 수 있습니다. 또한, 사용자들에게 개인화된 쇼핑 경험을 제공하는데 활용될 수 있습니다.

## 실습: Prompt Engineering

- zero shot prompting

다음의 문장을 영어로 번역하세요: '안녕하세요, 잘 지내세요?'

우리 태양계에서 가장 큰 행성은 무엇인가요?

- one shot prompting

다음은 빨간 사과에 대한 설명입니다. "빨간 사과는 맛있고, 영양분이 풍부하다."  
이와 유사한 방식으로 파란 바나나에 대해 설명하세요.

다음은 김영희에 대한 자기소개입니다: "안녕하세요. 저는 김영희입니다. 서울에서 프로그래머로 일하고 있습니다."  
이와 유사한 형식으로 박철수에 대한 자기소개를 만들어주세요.

- Few shot prompting

영화 제목을 짧게 요약하세요.  
영화 '타이타닉': 큰 배가 침몰하는 이야기.  
영화 '어벤저스': 슈퍼히어로들이 세계를 구하는 이야기.  
영화 '기생충'에 대해 짧게 요약해주세요.

다음은 몇몇 동물들의 특징을 설명한 것입니다.  
고양이: 독립적이며, 발바닥에 무릎이 있다.  
강아지: 충성스럽고, 꼬리를 흔들 때 행복하다.  
토끼에 대한 특징을 설명해주세요.

- Few shot COT(chain-of-thought)

만약 한 상자에 24개의 사과가 있고, 당신이 그 중 3분의 2를 먹었다면, 몇 개의 사과가 남아있을까요?" 단계별로 생각하세요

- 위 예제들을 변형해 가며 ChatGPT, Gemini, CLOVA X 에서 비교
  - COLVA X : <https://clova-x.naver.com/>
  - Gemini : <https://bard.google.com/chat>

# 생성형 Image Model



## 소목차

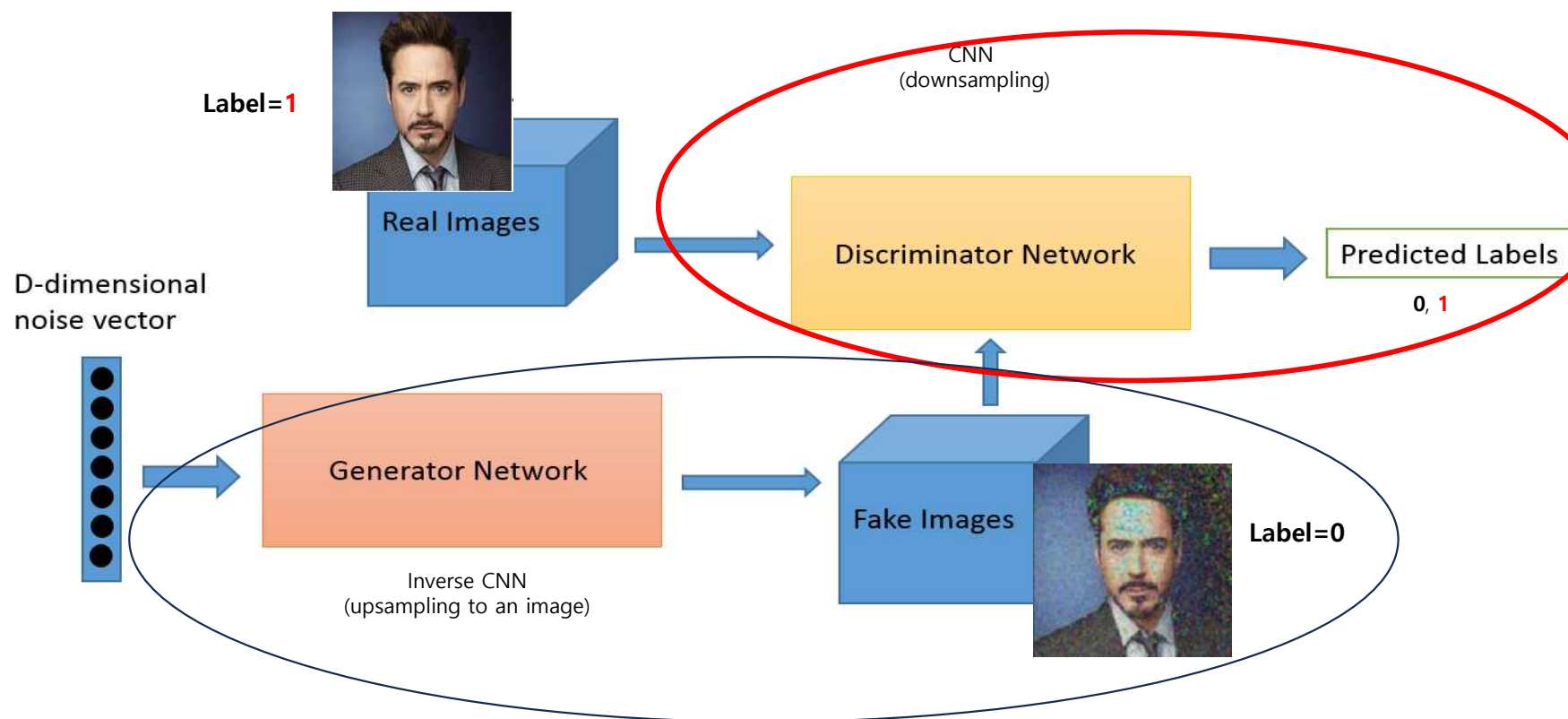
---

- GAN
- Diffusion Model
- DALL-E2
- Multi-modal Model

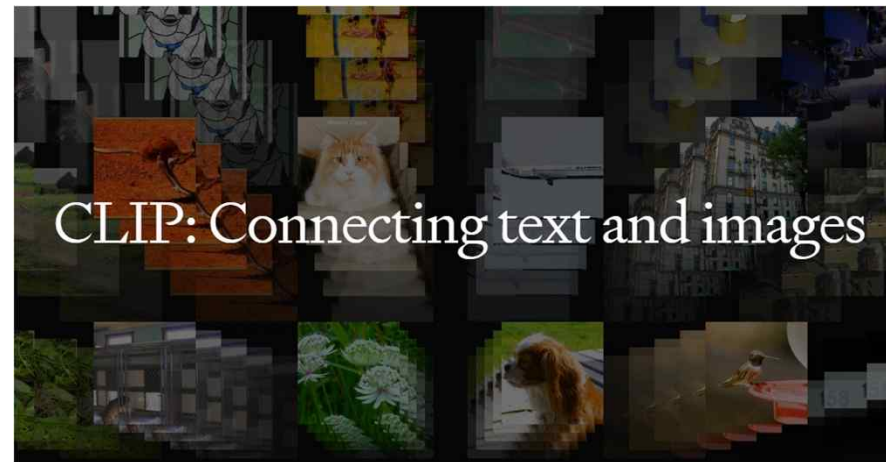
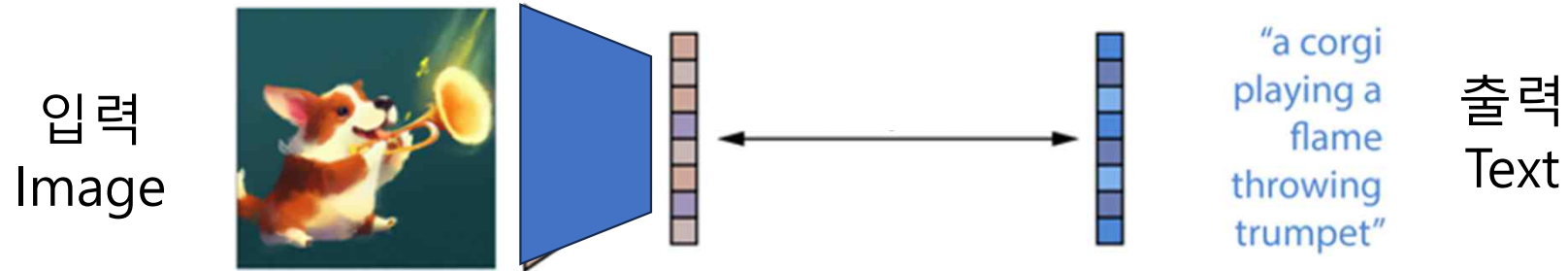
## GAN (Generative Adversarial Network)

- “the most interesting idea in the last 10 years in ML” – Yann LeCun
- 2014 Ian Goodfellow 가 최초 제안
- Computer 가 이미지, 인간의 목소리, 악기소리 등을 실제와 같이 생성
- 위조를 담당하는 Generator(생성자) 와 위조를 판별하는 Discriminator (판별자)의 두개의 Deep Neural Network 으로 구성

## GAN process

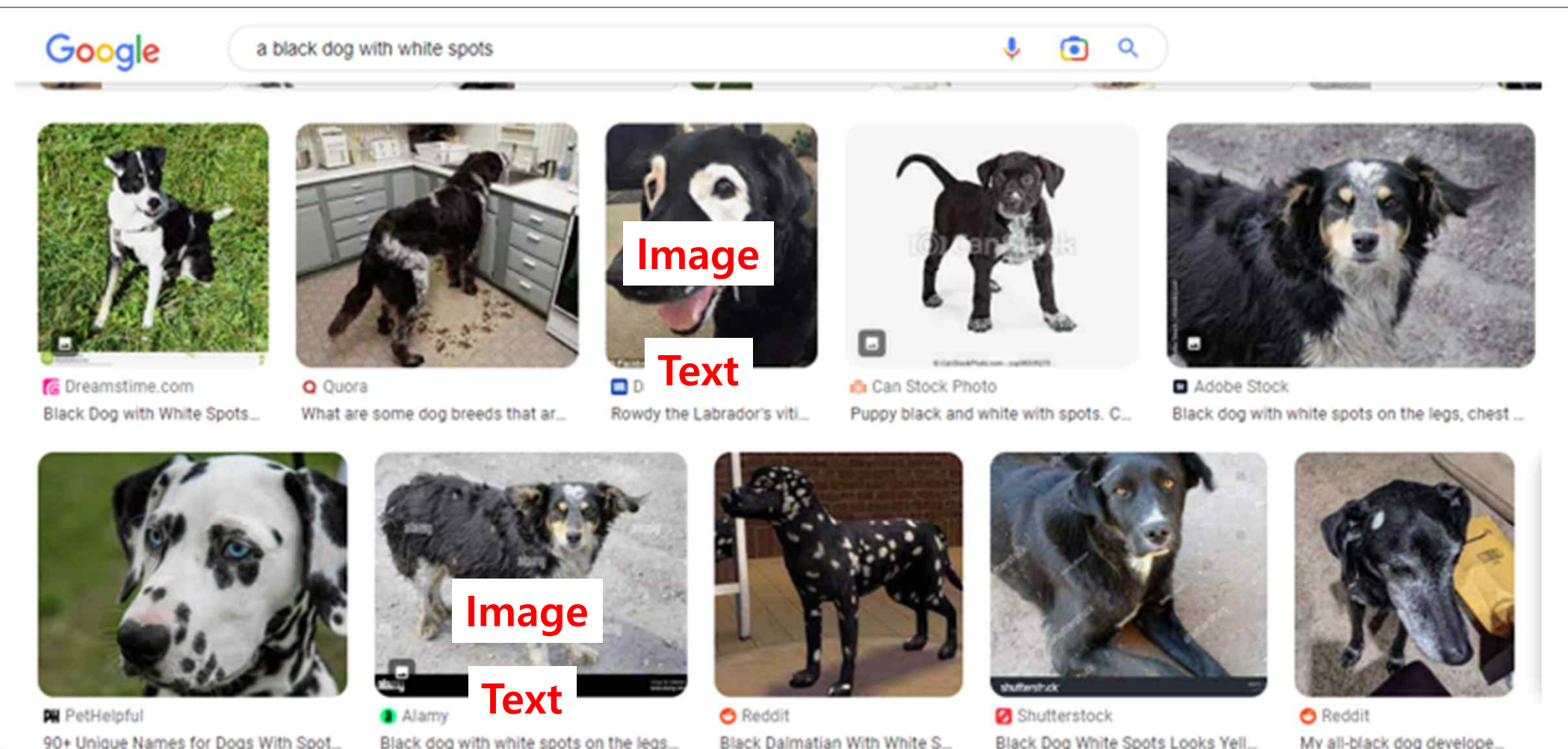


## Image → Text 생성 AI (OpenAI CLIP)

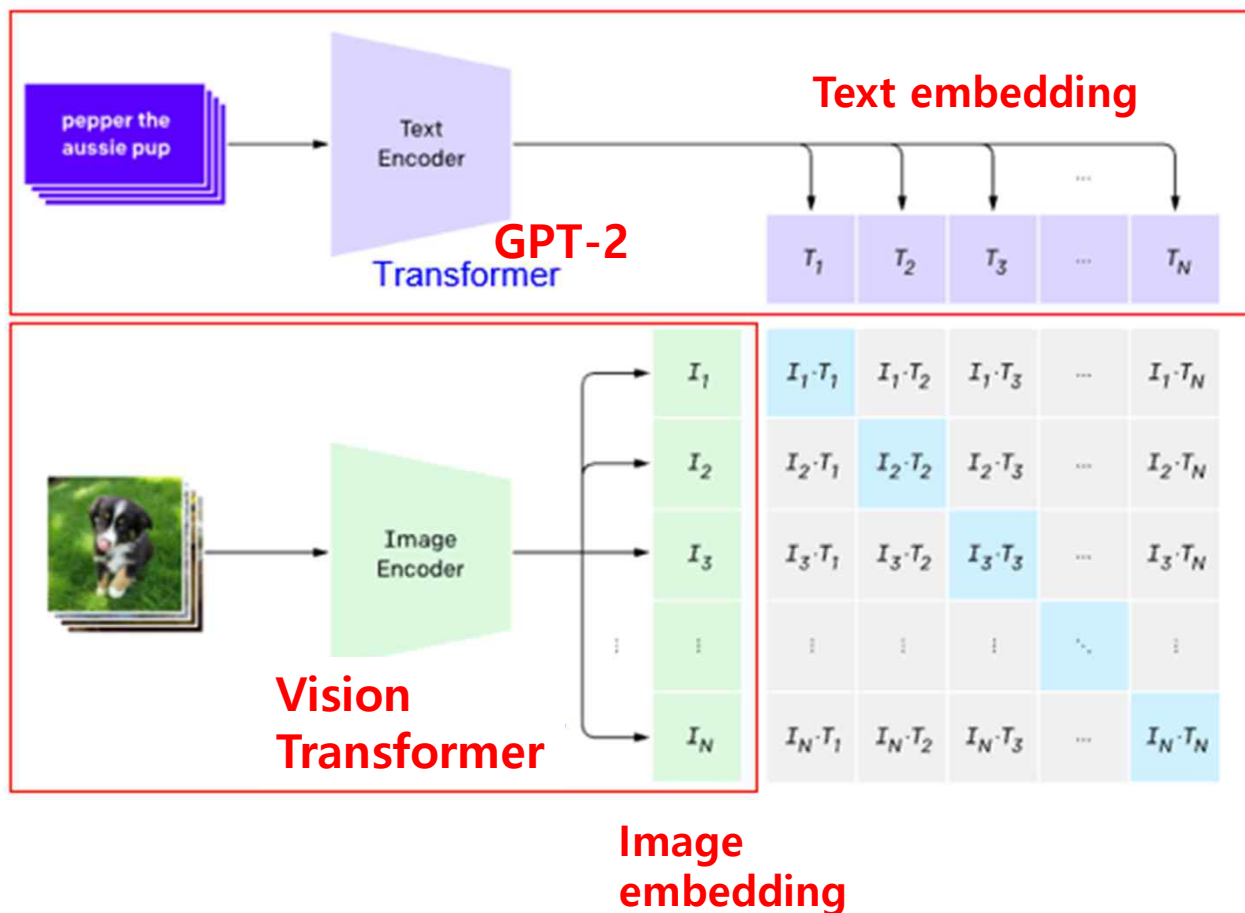




## Data example



## Contrastive Pre-Training



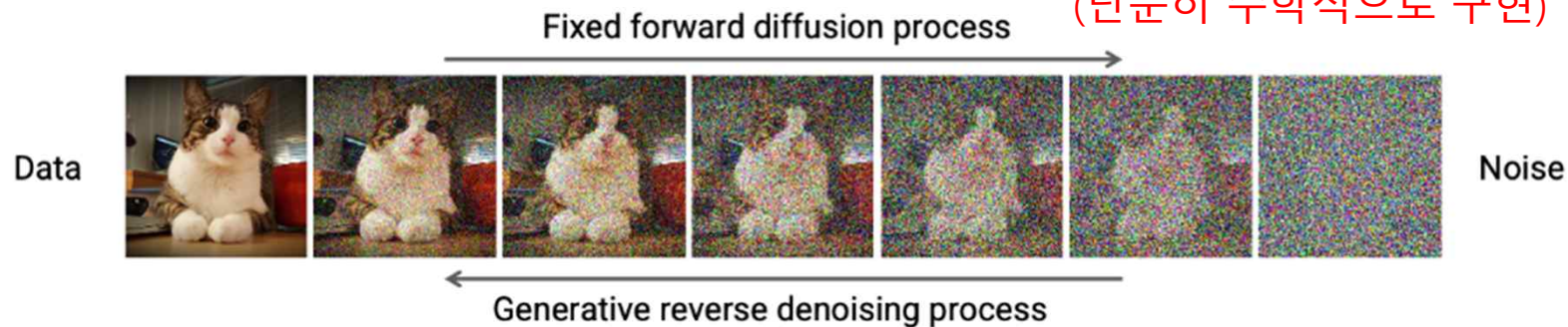
- N개의 이미지, 텍스트를 각각 encoder를 통과시켜 embedding vector 계산 (N: batch 사이즈)
- $N \times N$  matrix의 **대각선 부분이 positive pair**, 그 외 부분은 모두 **negative pair**  
(N개의 positive pair,  $N^2 - N$ 개의 negative pair)
- **Positive pair의 cosine 유사도는 maximize** 하고, **negative pair의 cosine 유사도는 minimize** 하도록 train

## Diffusion Model ?

- 확산 모델은 점진적인 노이즈 프로세스를 반대로 하여 데이터를 생성하는 방법을 학습
- GAN 수준의 High Quality + GAN 보다 뛰어난 Diversity 구현

이미지에 노이즈를 점진적으로 추가

(단순히 수학적으로 구현)



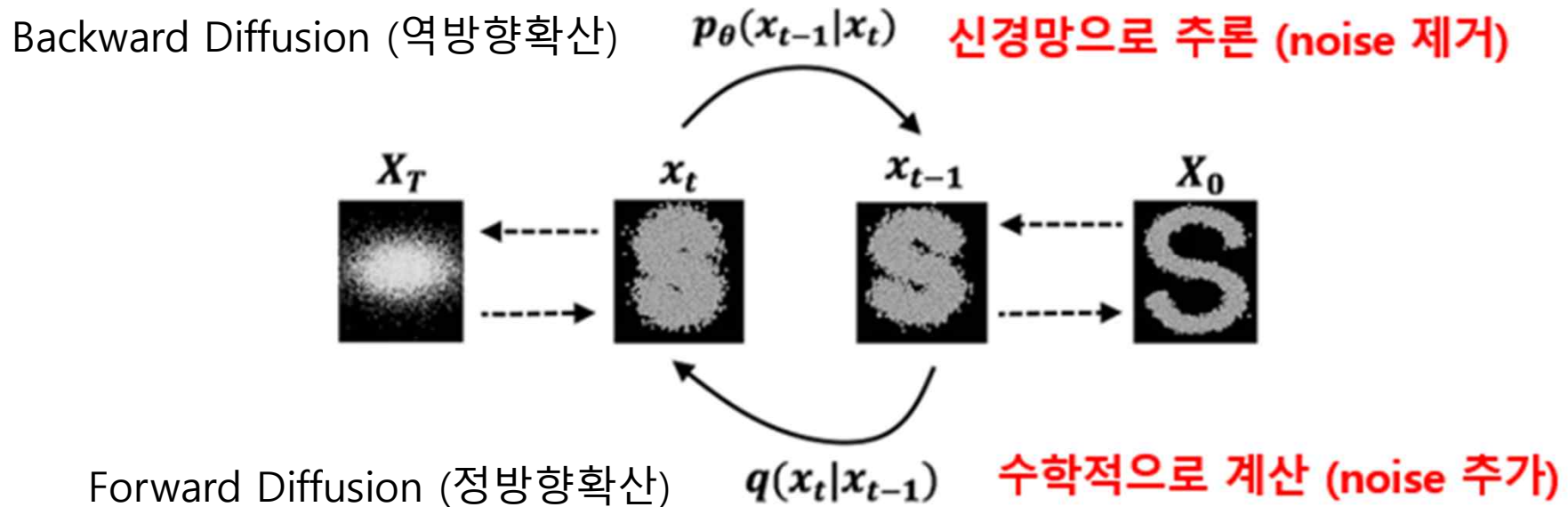
시간 단계에 걸쳐 점진적으로 노이즈를 제거

(신경망 학습)

## Diffusion model - Image 생성

- 두 단계로 작동

1. 데이터가 점차적으로 확산되는 방식을 시뮬레이션하여 잠재 변수를 생성
2. 잠재 변수를 사용하여 데이터를 재구성



## GAN vs. Diffusion model

특징	GAN	Diffusion model
기본 아이디어	생성자와 판별자의 경쟁	노이즈로부터 데이터를 생성하는 역변환 모델
학습 과정	생성자와 판별자를 동시에 학습	역변환 모델을 학습 (U-Net 이용)
장점	고품질의 생성물을 생성할 수 있음	학습이 안정적임
단점	<ul style="list-style-type: none"> <li>- Mode collapse가 일어나기 쉽고 학습이 어려움</li> <li>- 생성물의 다양성이 떨어짐</li> </ul>	생성물이 현실적이지 않을 수 있음

- GAN은 noise로 부터 한번에 image를 만들지만 Diffusion model은 noise로 부터 step t 만큼 조금씩 noise를 빼며 image를 만든다.

## What is DALL-E ?

- 2020. 9.28에 OpenAI에서 발표한 생성형 AI 모델이며 GPT-3와 같은 transformer language model : 텍스트 ➔ 고품질 이미지 생성
- 동물과 물체의 의인화된 버전 생성, 관련 없는 개념을 그럴듯한 방식으로 결합, 텍스트 렌더링, 기존 이미지에 변형 적용 등 다양한 기능
- 기존에 존재 하는 이미지가 아닌 완전히 새로운 이미지를 생성  
➔ 높은 Zero-shot performance  
(ex. 학습 데이터에 "자전거", "불" 포함 ➔ "불타는 자전거" 이미지 생성)

- DALL-E : 120 억개의 매개변수로 이루어진 Transformer 기반의 GPT-3 확장 형태  
→ 2.5억개의 (텍스트, 이미지) 쌍으로 학습
- Computer Vision 과 NLP 기술을 결합하여 만든 Text to Image Task를 Auto-regressive 하게 모델링
- DALL-E2 : DALL-E의 개선된 version. CLIP + Diffusion

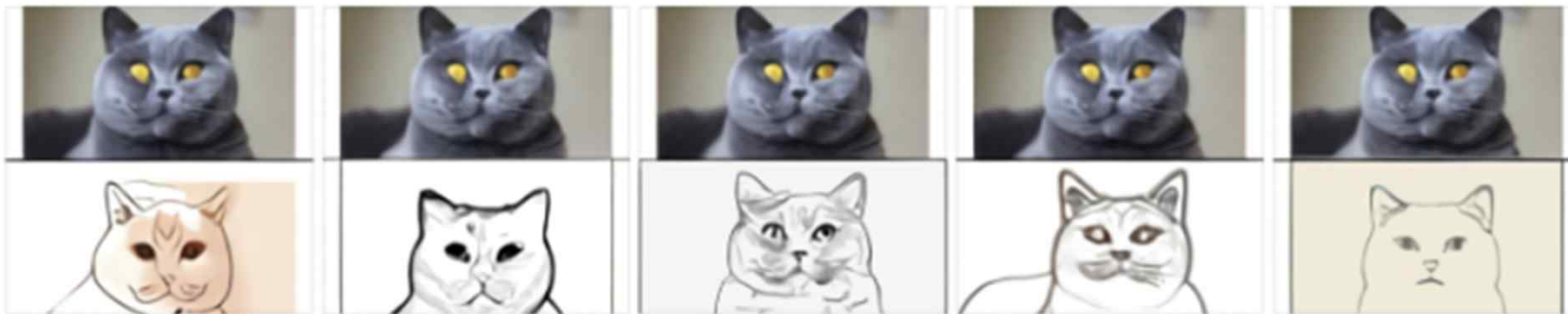


## DALL-E 생성 이미지

- Prompt: an arm chair in the shape of and avocado (의자 모양의 아보카도)

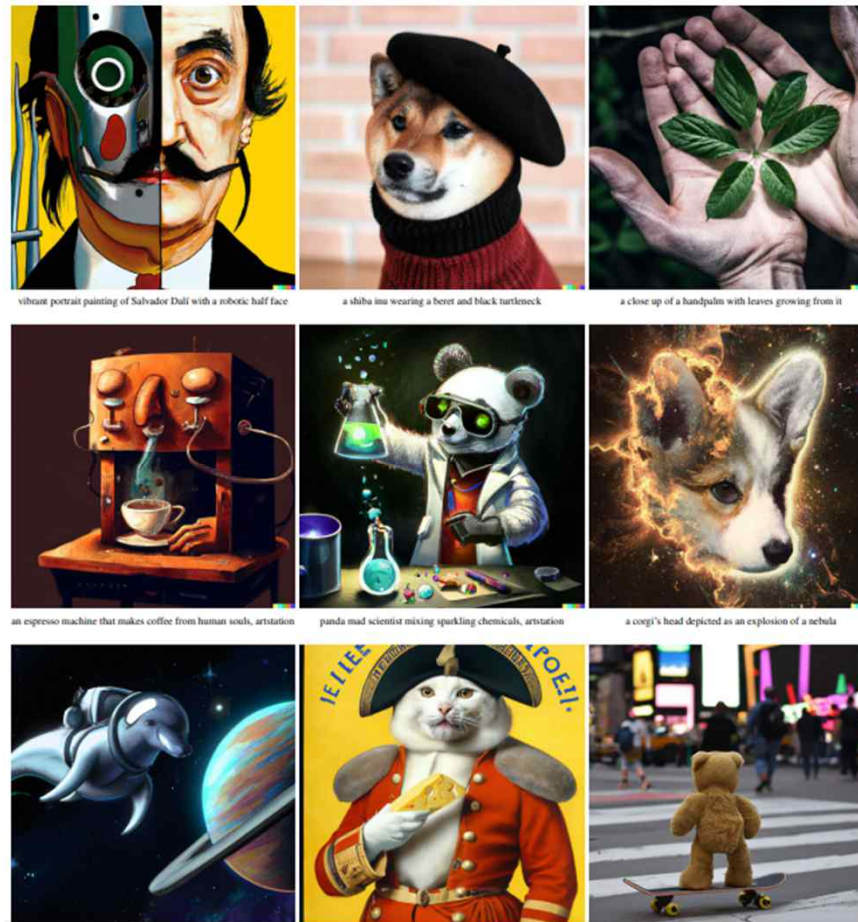


- Prompt: the exact same cat on the top as a sketch on the bottom (하단의 스케치와 정확히 같은 상단의 고양이)





## DALL-E2 생성 이미지



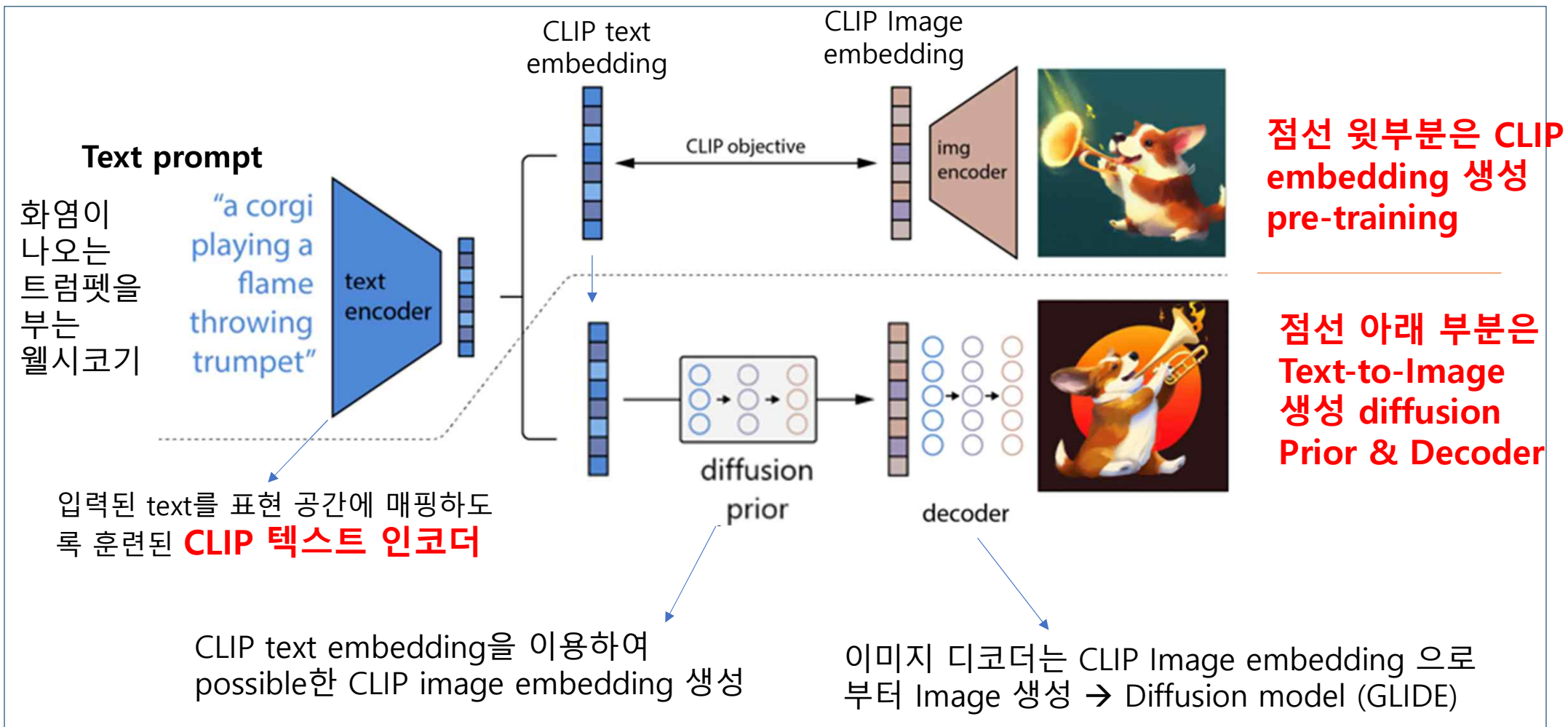
## DALL-E3 생성 이미지



Larry의 집을 그려줘 ➡

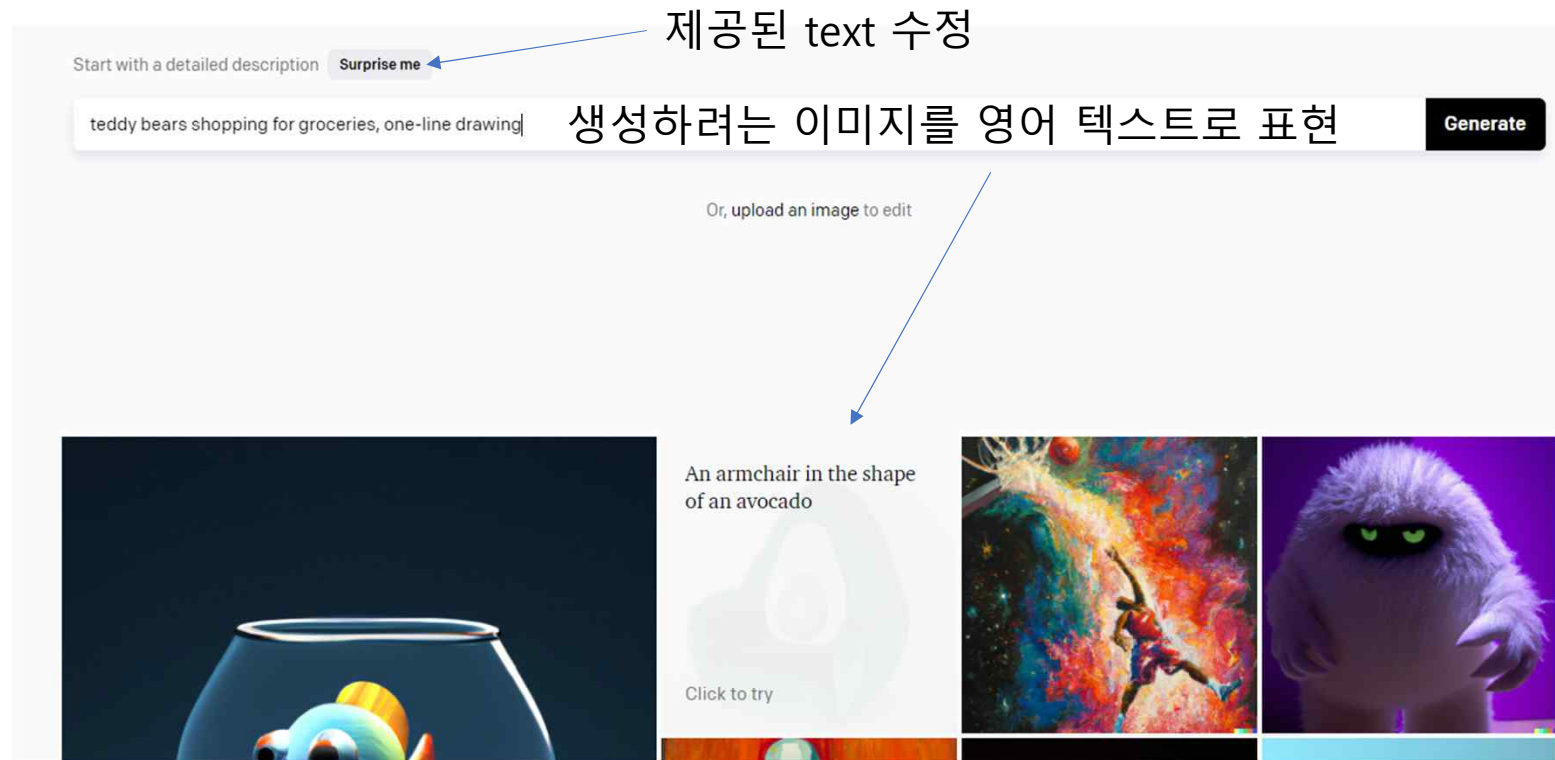


## DALL-E2 Structure



## 실습: openAI DALL-E2

- <https://labs.openai.com/>





## 기능 1 : (영문) Text에서 image 생성

Edit the detailed description

Surprise me

Upload →

Dancing BTS on Everest Mountain

Generate

# Dancing BTS on Everest Mountain



## LMM(Large Multimodal Model) 이란 ?

- Large Language Models (LLMs)의 기능을 확장하여 시각적 이해를 포함한 다중 감각 기술을 통합한 모델
  - 언어 처리 능력에 더하여 이미지, 비디오 등의 시각적 데이터를 이해하고 처리할 수 있는 능력을 갖추므로써 더 강력한 범용 인공지능을 실현
- OpenAI의 GPT-4가 대표적인 LMM

# GPT-4

- GPT-4는 다양한 벤치마크에서 인간 수준의 성능을 보여주는 대규모 다중 모드 모델(이미지 및 텍스트를 입력, 텍스트 출력)
- 상위 10% 정도의 점수로 변호사 시험을 통과 (GPT-3.5의 점수는 하위 10% 정도)
- <https://openai.com/research/gpt-4> 참조

## GPT-4 Examples

### Image-text Pairs

- Prompt:  
메뉴에 적힌 가격에 따라 테이블 위의 맥주에 얼마를 지불해야 합니까?
- Response:  
메뉴에 따르면 마그나 맥주 가격은 6달러입니다. 테이블 위에 마그나 맥주 두 개가 있으므로 맥주의 총 비용은 12달러가 됩니다.

How much should I pay for the beer on the table according to the price on the menu?



Here is a photo of the menu.

Wine & Beers			
<b>White</b>			
Pinot Grigio, Gavi, Glera	1545	<b>Red</b>	
Chardonnay, Sicily	1545	Corvina, Sangiovese, Lombard	1500
Bucca Verdisio, D.O.C.G. 2018	1559	<b>Sparkling</b>	
		Prosecco, Glera	1640
<b>Red</b>			
Sauvignon, Cabernet, Toscana	1334	<b>Beers</b>	
Cal di Sesto, Bard, Toscana	1040	Magda	5
Sangiovese, Toscana	1140	Alga	5
		Acqua	5
		Acqua	6



# Image-text Pairs

- Prompt:  
내가 낸 세금은 얼마였나요?  
정보는 어디에 있나요?
- Response:  
첫 번째 영수증에는 세금으로 \$3.72를 지불 하셨습니다. 정보는 영수증 하단에 있습니다.  
두 번째 영수증에서는 세금으로 \$42.23를 지불했습니다. 해당 정보는 영수증 하단의 " 판매세" 라인 아래에 있습니다.  
세 번째 영수증에는 세금으로 \$4.58를 지불 하셨습니다. 정보는 영수증 하단의 "TAX" 라인 아래에 있습니다.  
총 세 가지 영수증에 걸쳐 세금으로 \$50.53를 납부하셨습니다.

How much did I pay for tax? Prompt Continued  
Where is the information?



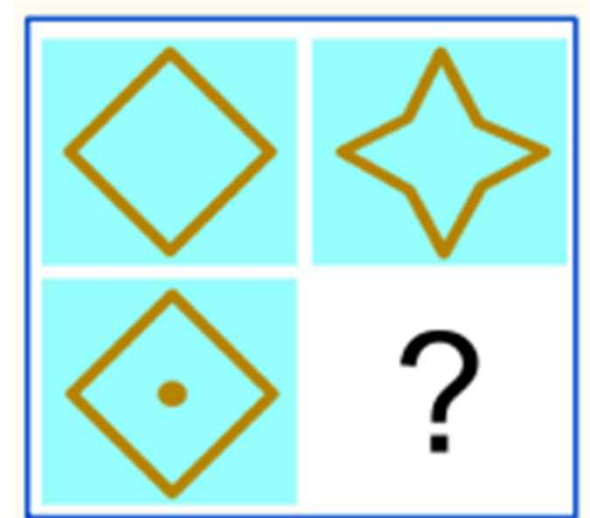
## Image-text Pairs

- Prompt:

다음은 2x2 행렬로 배열된 세 개의 하위 이미지입니다. 먼저 첫 번째 열의 왼쪽 상단과 왼쪽 하단에 있는 두 이미지를 살펴보세요. 그런 다음 왼쪽 두 이미지를 기반으로 첫 번째 열의 패턴을 찾습니다. 다음으로, 찾은 패턴과 오른쪽 상단의 이미지를 활용하여 누락된 도형을 추론해 보세요. 마지막으로 누락된 이미지는 어떤 모습일까요?

- Response:

첫 번째 열의 패턴은 왼쪽 하단 이미지의 모양이 왼쪽 상단 이미지의 모양 중앙에 점이 있다는 것입니다. 따라서 오른쪽 하단의 누락된 이미지는 중앙에 점이 있는 별이 됩니다. 이는 오른쪽 상단의 이미지와 비슷하지만 중앙에 점이 있는 것입니다.



## Image-text Pairs

- Prompt: 이미지에 있는 사과의 개수를 셉니다.
- Response: 이미지에는 12개의 사과가 있습니다.

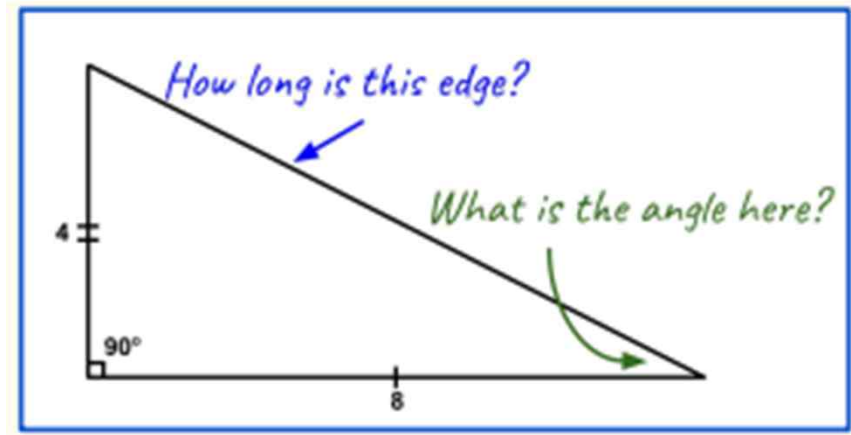


- Prompt:  
이미지 속 인물은 누구이며, 무엇을 하고 있나요?.
- Response:  
사진 속 인물은 현 미국 대통령 조 바이든입니다. 그는 2023년 일본 히로시마에서 열리는 G7 정상회의에서 연설을 할 예정으로 연단에 서 있습니다.



## Visual Referring

- Prompt: 다음 그림의 질문에 대해 보세요.  
중간 단계를 제공하세요.



- Response:  
파란색 화살표 모서리의 길이를 찾으려면 직각 삼각형이 있으므로 피타고라스 정리를 사용할 수 있습니다. 정리에 따르면  $a^2 + b^2 = c^2$ 입니다. 여기서 a와 b는 직각삼각형의 두 다리이고 c는 빗변입니다. 우리가 가지고 있는 값을 대입하면 다음과 같습니다:

$$4^2 + 8^2 = c^2 \quad 16 + 64 = c^2 \quad 80 = c^2 \quad \sqrt{80} = c \quad 8.94 \approx c$$

따라서 파란색 화살표 가장자리의 길이는 대략 다음과 같습니다. 8.94.

녹색 화살표의 각도를 찾으려면 각도의 반대쪽과 인접한 쪽이 있으므로 접선 기능을 사용할 수 있습니다. 탄젠트 함수는  $\tan\theta = \text{반대쪽}/\text{인접함}$ 을 나타냅니다. 우리가 가지고 있는 값을 연결하면 다음과 같은 결과를 얻습니다.  $\tan\theta = 4/8 \quad \tan\theta = 1/2 \quad \theta = \tan^{-1}(1/2) \quad \theta \approx 26.57^\circ$

따라서 녹색 화살표의 각도는 약  $26.57^\circ$ 입니다.