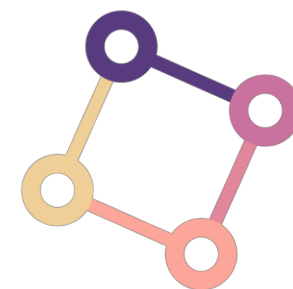


자기지도학습 ⊕ 언어 모델을 통한 대규모 사전학습 모델

Self-Supervised Learning and Large-Scale Pre-Trained Models

주재걸 교수

KAIST 김재철AI대학원

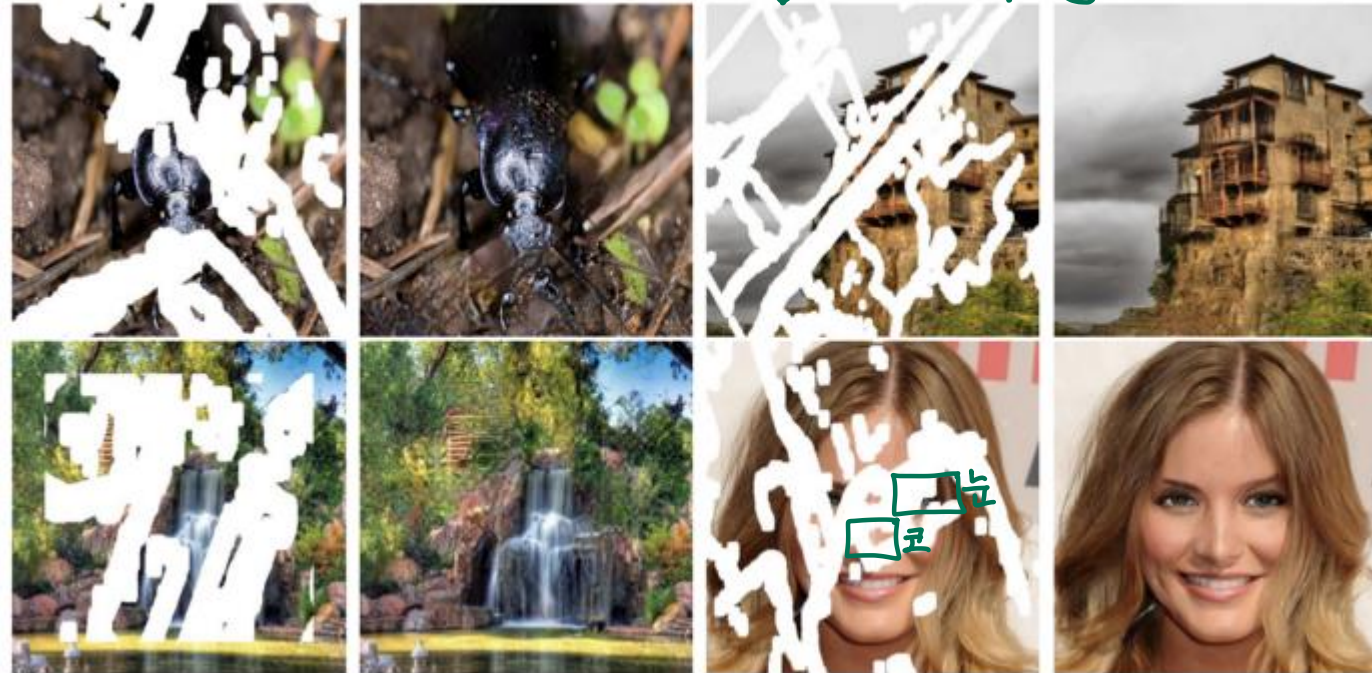


DAVIAN

Data and Visual Analytics Lab

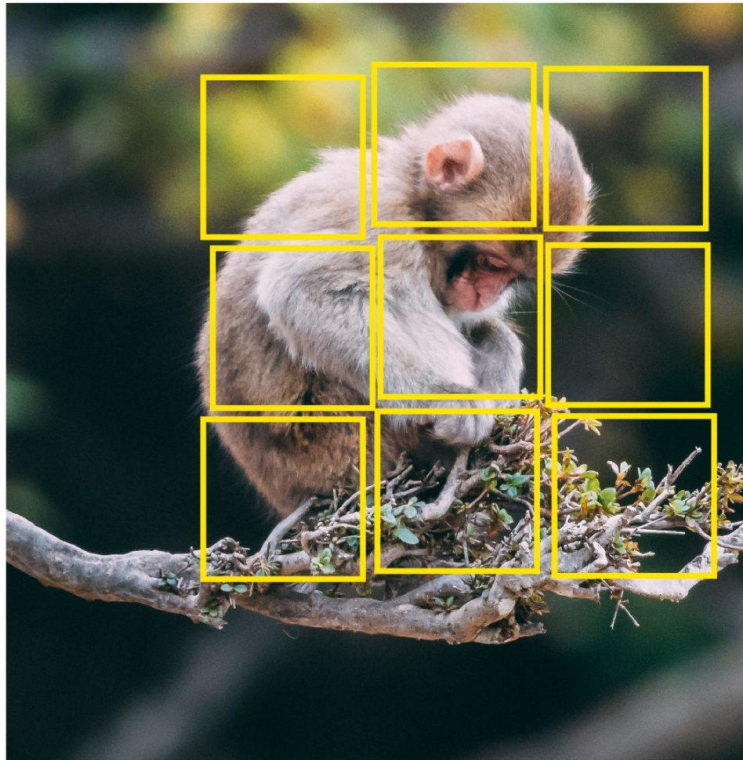
What is Self-Supervised Learning? 자기지도 학습?

- Given unlabeled data, hide part of the data and train the model so that it can predict such a hidden part of data, given the remaining data.
- Image in-painting example:

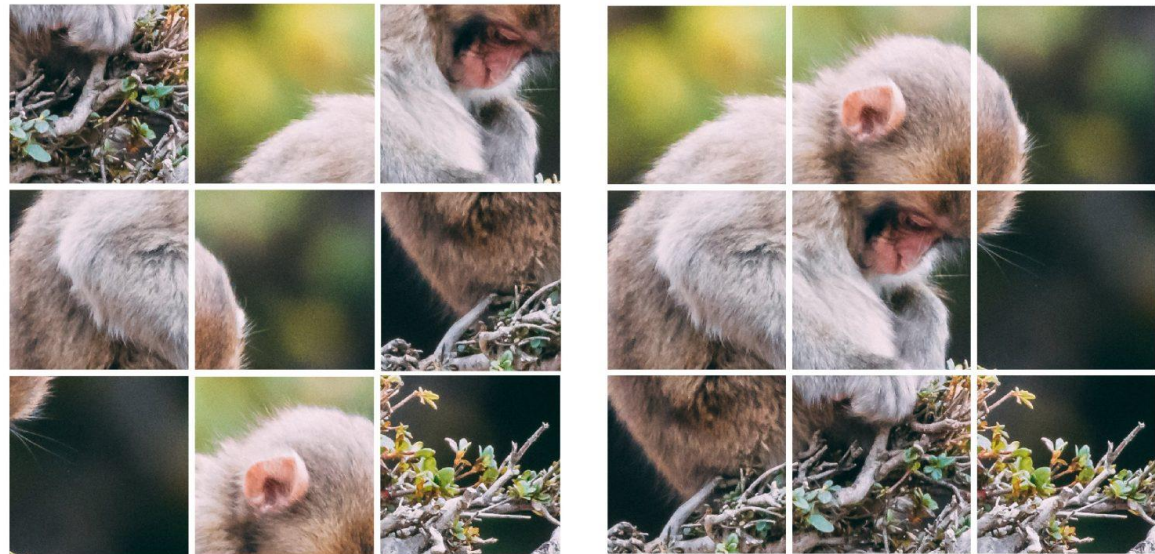


What is Self-Supervised Learning?

- Given unlabeled data, hide part of the data and train the model so that it can predict such a hidden part of data, given the remaining data.
- Solving a zigsaw puzzle:

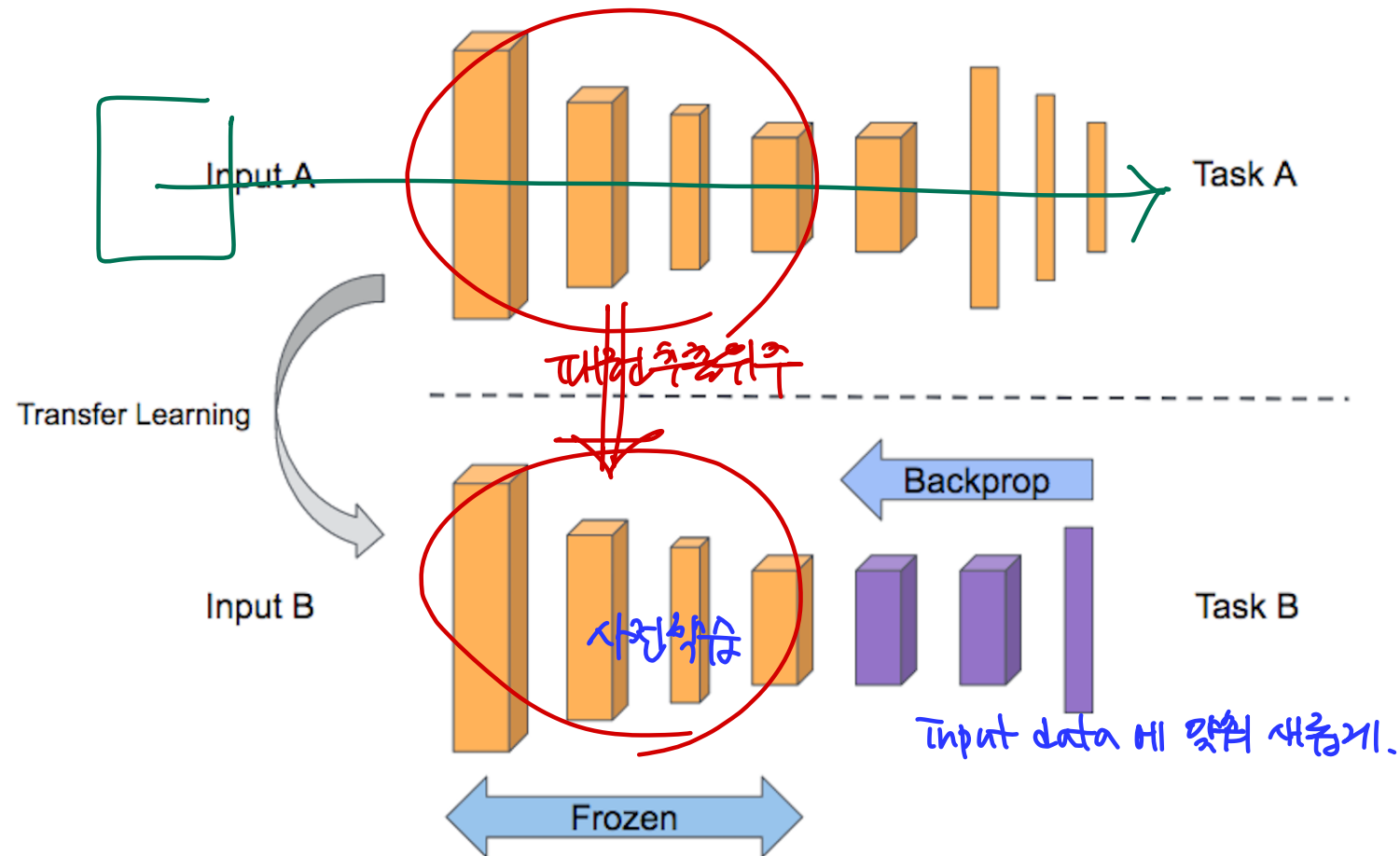


익숙파uzzle



Transfer Learning from Self-Supervised Pre-trained Model

- Pre-trained models using a particular self-supervised learning can be fine-tuned to improve the accuracy of a given target task.



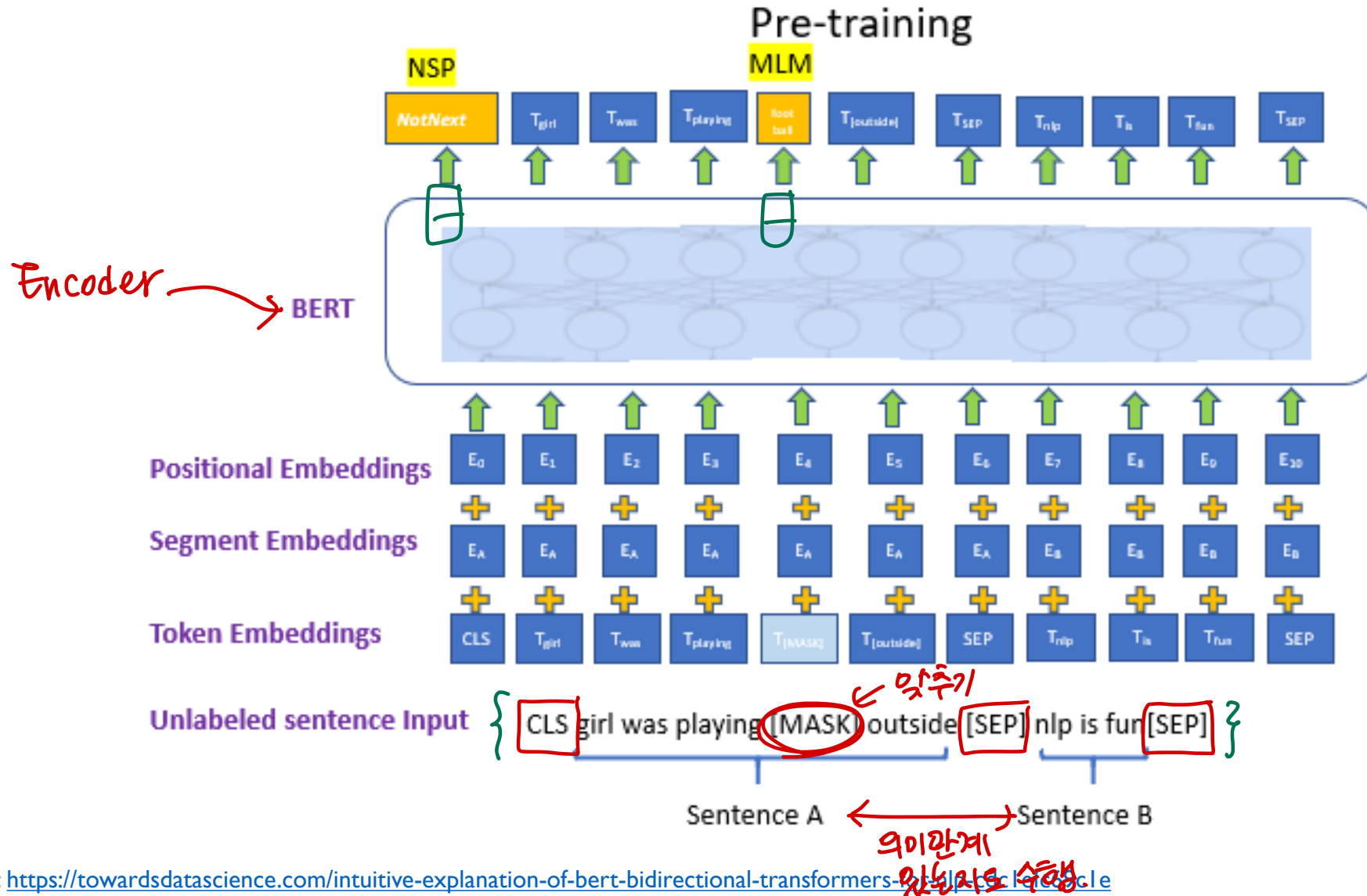
BERT

- BERT: Pre-training of Deep **Bidirectional** Transformers for Language Understanding

- Learn through masked language modeling (MLM) and next-sentence prediction (NSP) tasks
- Use large-scale data and large-scale model

자바 21도 학습 .

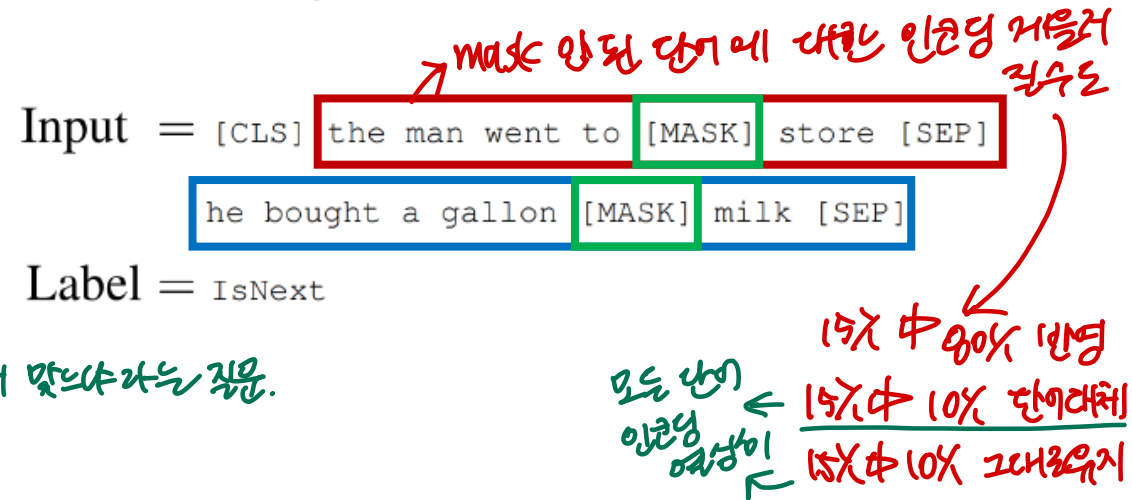
Overview of BERT



Pre-Training Tasks of BERT

- Masked Language Model (MLM)

- Mask some percentage of the input tokens at random, and then predict those masked tokens.
- 15% of the words to predict
 - 80% of the time, replace with [MASK]
 - 10% of the time, replace with a random word
 - 10% of the time, keep the sentence as same



- Next Sentence Prediction (NSP)

- Predict whether **Sentence B** is an actual sentence that proceeds **Sentence A**, or a random sentence

Masked Language Modeling (MLM) Task

- Mask out $k\%$ of the input words, and then predict the masked words.

- BERT uses $k=15\%$

- Too little masking : Too expensive to train.

- Too much masking : Not enough to capture the given context. → *배고지만 머쓱어.*

the man went to the [MASK] to buy a [MASK] of milk

store gallon

↑ ↑

- Out of $k=15\%$ selected tokens,

- For 80% of the time, replace them with [MASK] token

- went to the store → went to the [MASK]

- For 10% of the time, replace them with a random word

- e.g., went to the store → went to the running

- For 10% of the time, keep them as it is,

- e.g., went to the store → went to the store

Next Sentence Prediction (NSP) Task

- To learn the relationships among sentences, predict whether Sentence B is an actual sentence that follows Sentence A, or a random, irrelevant sentence

Input = { [CLS] the man went to [MASK] store [SEP]
he bought a gallon [MASK] milk [SEP] }

Label = IsNext

Handwritten notes:
- Red arrow from [CLS] to [MASK] in the first sentence: *연속된 두 개 문장인지*
- Green circle 1 around [CLS]
- Green circle 2 around the first [MASK]
- Red curly brace around the second sentence

Input = [CLS] the man [MASK] to the store [SEP]
penguin [MASK] are flight ##less birds [SEP]

Label = NotNext

Further Details of BERT

1. Model Architecture

- BERT BASE: $L = 12$, $H = 768$, $A = 12$
- BERT LARGE: $L = 24$, $H = 1024$, $A = 16$

layer 7
Hidden state

head 11A (set)

2. Input Representation

- WordPiece embeddings (30,000 WordPiece)
- Learned positional embedding
- [CLS] – Classification embedding
- Packed sentence embedding [SEP]
- Segment Embedding

3. Pre-training Tasks

- Masked LM
- Next Sentence Prediction

Segment Embedding and Position Embedding

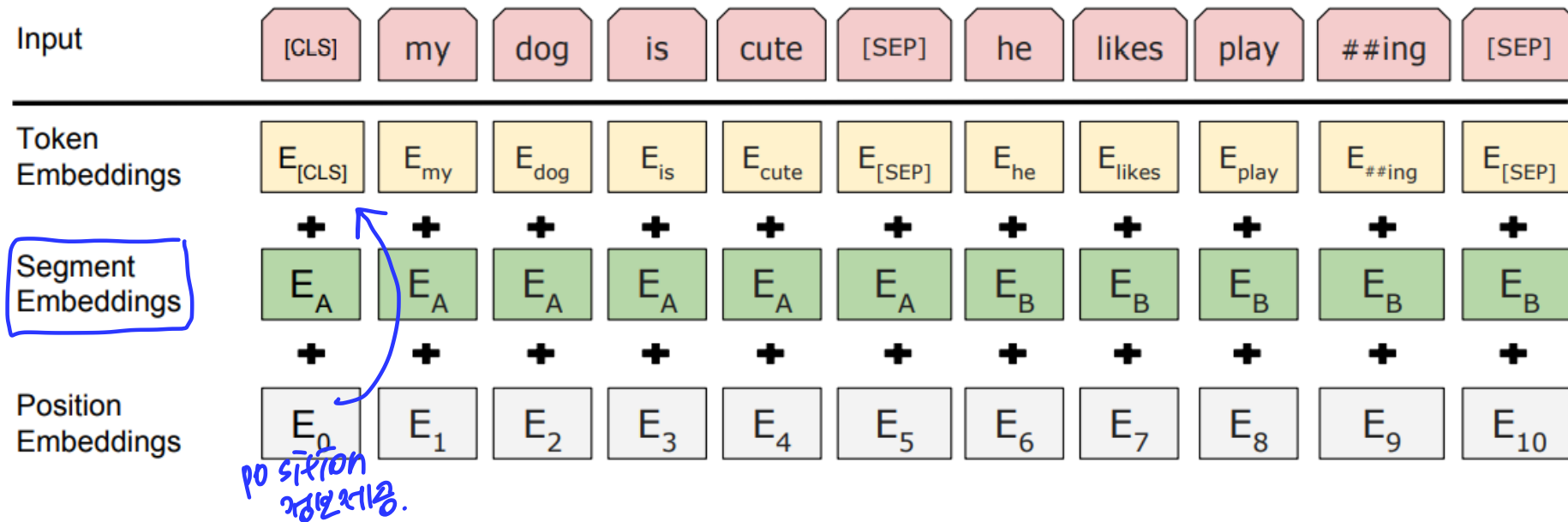
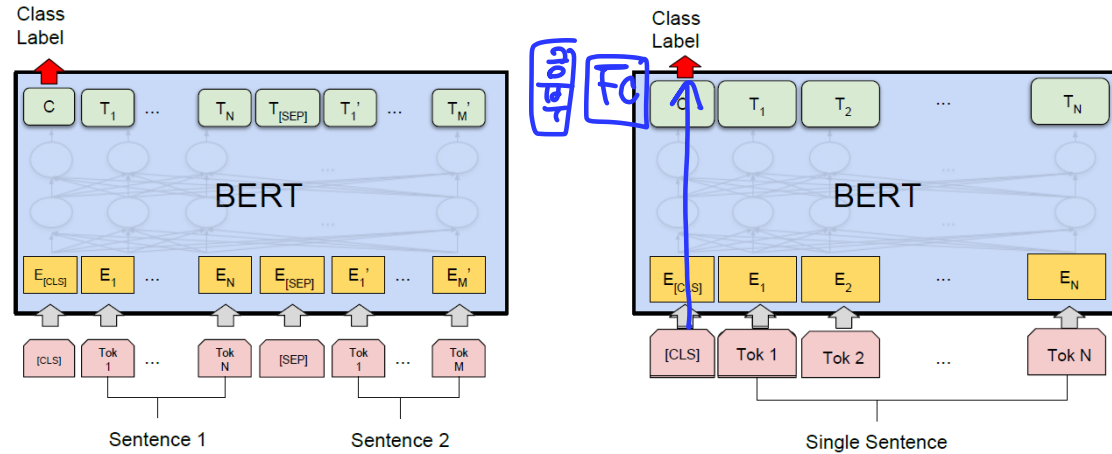


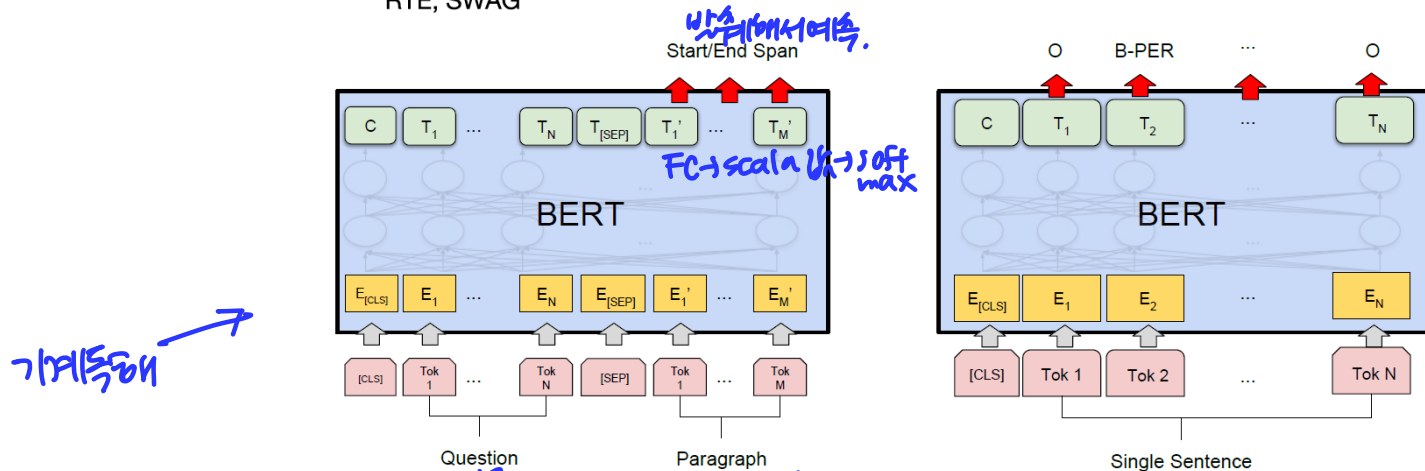
Figure 2: BERT input representation. The input embeddings is the sum of the token embeddings, the segmentation embeddings and the position embeddings.

Various Fine-Tuning Approaches



(a) Sentence Pair Classification Tasks:
MNLI, QQP, QNLI, STS-B, MRPC,
RTE, SWAG

(b) Single Sentence Classification Tasks:
SST-2, CoLA



(c) Question Answering Tasks:
SQuAD v1.1

(d) Single Sentence Tagging Tasks:
CoNLL-2003 NER

BERT: GLUE Benchmark Results

System	MNLI-(m/mm) 392k	QQP 363k	QNLI 108k	SST-2 67k	CoLA 8.5k	STS-B 5.7k	MRPC 3.5k	RTE 2.5k	Average
Pre-OpenAI SOTA	80.6/80.1	66.1	82.3	93.2	35.0	81.0	86.0	61.7	74.0
BiLSTM+ELMo+Attn	76.4/76.1	64.8	79.9	90.4	36.0	73.3	84.9	56.8	71.0
OpenAI GPT	82.1/81.4	70.3	88.1	91.3	45.4	80.0	82.3	56.0	75.2
BERT _{BASE}	84.6/83.4	71.2	90.1	93.5	52.1	85.8	88.9	66.4	79.6
BERT _{LARGE}	86.7/85.9	72.1	91.1	94.9	60.5	86.5	89.3	70.1	81.9

• MultiNLI

- Premise: Hills and mountains are especially sanctified in Jainism.
- Hypothesis: Jainism hates nature.
- Label: Contradiction

어제 최모양 1명이 사망했다
어제 최모가 사망했다 (참이면)

• CoLa

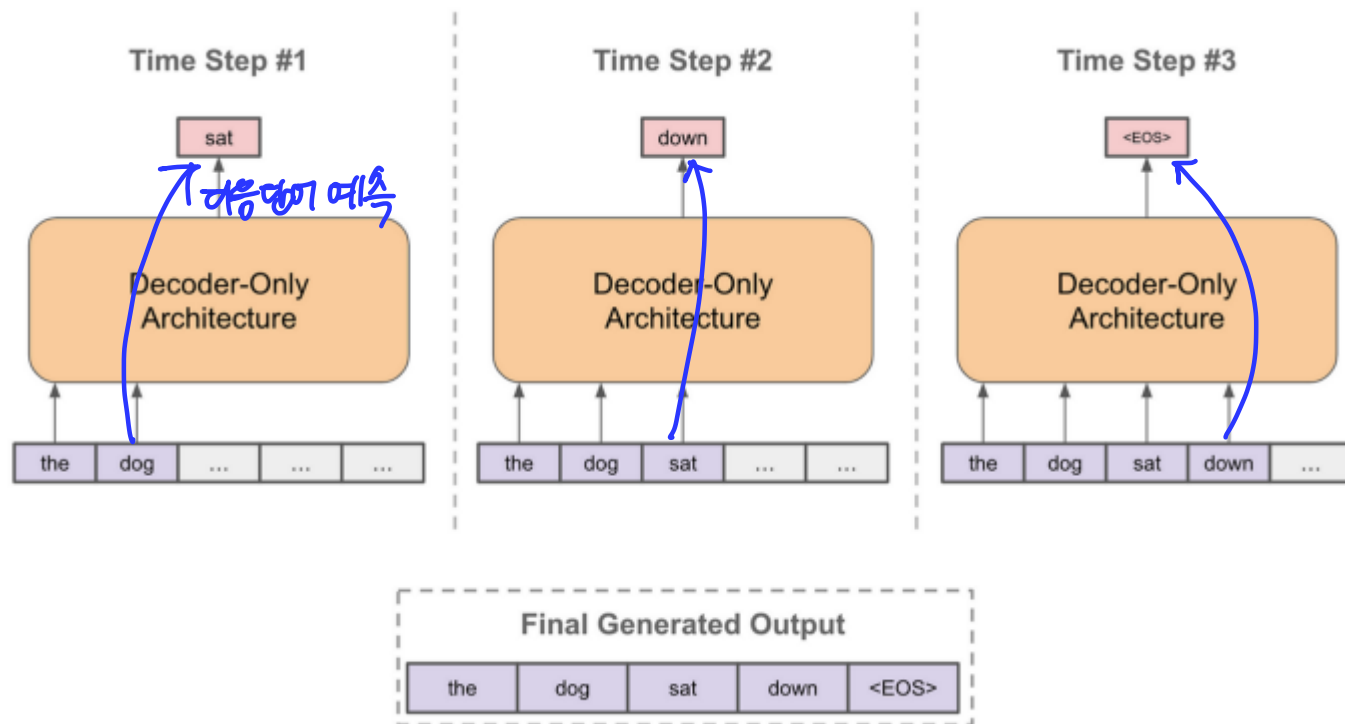
- Sentence: The wagon rumbled sown the road.
Label: Acceptable
- Sentence: The car honked down the road.
Label: Unacceptable

GPT-1/2/3: Generative Pre-Trained Transformer

- Generative Pre-Training Task

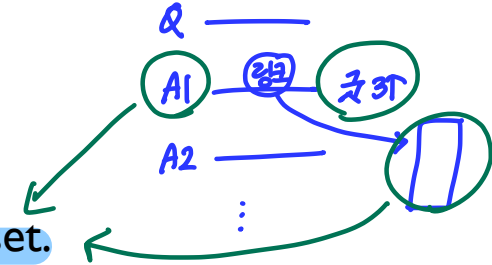
- In other words, this task is called Language Modeling.
- From another perspective, this task is called an Auto-Regressive Model, in a sense that the predicted output at the current time step will be used as an input at the next time step.

Decoder 사용!



GPT-2: Language Models are Unsupervised Multi-task Learners

- Just a really big transformer-based language model 모델 사이즈 큼.
 - Trained on 40GB of text ↓
 - A large amount of efforts have been put to secure a high-quality dataset.
 - Take webpages from reddit links with at least 3 karma (up-vote)
 - Language models can perform down-stream tasks in a zero-shot setting without any parameter or architecture modification



SYSTEM PROMPT
(HUMAN-WRITTEN)

prompt

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION
(MACHINE-WRITTEN, 10 TRIES)

output
completion

The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.

Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns.

While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic."

GPT-2: Zero-Shot Summarization

goal task

- CNN and Daily Main dataset
 - Add text TL;DR: after the article and generate 100 tokens (TL;DR: Too long, didn't read)

Language Models are Unsupervised Multitask Learners

Article: Prehistoric man sketched an incredible array of prehistoric beasts on the rough limestone walls of a cave in modern day France 36,000 years ago. Now, with the help of cutting-edge technology, those works of art in the Chauvet-Pont-d'Arc Cave have been reproduced to create the biggest replica cave in the world. The manmade cavern named the Caverne du Pont-d'Arc has been built a few miles from the original site in Vallon-Pont-D'arc in Southern France and contains 1,000 painstakingly-reproduced drawings as well as around 450 bones and other features... Cavemen and women sketched an incredible array of prehistoric beasts on the rough limestone walls of a cave 36,000 years ago and now a replica has been created (pictured) ...

GPT-2: The original site in Vallon-Pont-D'arc in Southern France is a Unesco World Heritage site and is the oldest known and the best preserved cave decorated by man. The replica cave was built a few miles from the original site in Vallon-Pont-D'Arc in Southern France. The cave contains images of 14 different species of animals including woolly rhinoceros, mammoths, and big cats.

Reference: Cave mimics famous Caverne du Pont-d'Arc in France, the oldest cave decorated by man and the best preserved. The replica contains all 1,000 paintings which include 425 such as a woolly rhinoceros and mammoths. Minute details were copied using 3D modelling and anamorphic techniques, often used to shoot widescreen images. The modern cave also includes replica paw prints of bears, bones and details preserved in the original cave.

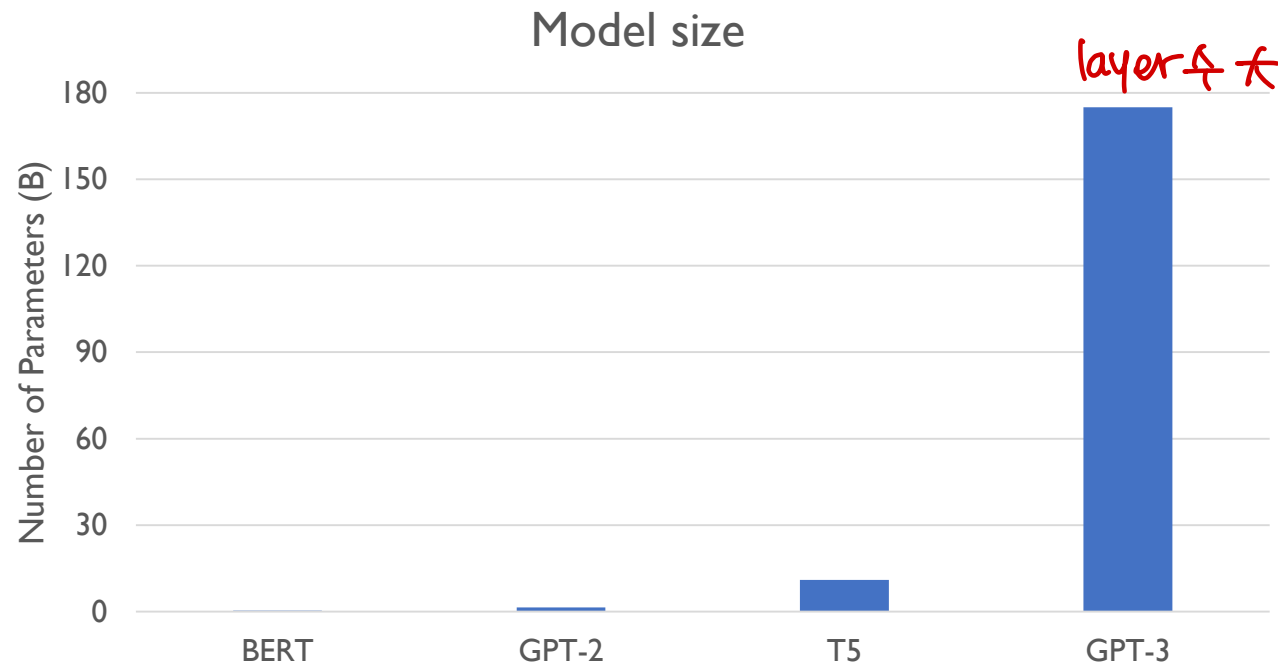
TL;DR ← 마지막 단어

	R-1	R-2	R-L	R-AVG
Bottom-Up Sum	41.22	18.68	38.34	32.75
Lede-3	40.38	17.66	36.62	31.55
Seq2Seq + Attn	31.33	11.81	28.83	23.99
GPT-2 TL;DR:	29.34	8.27	26.58	21.40
Random-3	28.78	8.63	25.52	20.98
GPT-2 no hint	21.58	4.03	19.47	15.03

Table 4. Summarization performance as measured by ROUGE F1 metrics on the CNN and Daily Mail dataset. Bottom-Up Sum is the SOTA model from (Gehrmann et al., 2018)

GPT-3: Language Models are Few-Shot Learners

- Scaling up language models greatly improves task-agnostic, few-shot performance
- An autoregressive language model with 175 billion parameters in the few-shot setting
한정 데이터가.
- 96 attention layers, batch size of 3.2M, 175B parameters



Few-Shot Learning Example of GPT-3

- **Prompt:** the prefix given to the model
- **Zero-shot:** Predict the answer given only a natural language description of the task
- **One-shot:** See a single example of the task in addition to the task description
- **Few-shot:** See a few examples of the task

```
1 Translate English to French: ← task description
2 cheese =>  ← prompt
```

Zero-shot

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← example
3 cheese =>  ← prompt
```

One-shot learning task

prompt tuning task
어떻게 할 것인가도 중요

```
1 Translate English to French: ← task description
2 sea otter => loutre de mer ← examples
3 peppermint => menthe poivrée ←
4 plush girafe => girafe peluche ←
5 cheese =>  ← prompt
```

Few-shot ← 성능 좋음.

Copilot: Program Auto-Completion based on GPT-3



프로그래밍 자동완성 코드

Don't fly solo

Enjoy a 60-day free trial then \$10/month or \$100/year per user

Start my free trial >

Explore docs

Your AI pair programmer

GitHub Copilot uses the OpenAI Codex to suggest code and entire functions in real-time, right from your editor.

Start my free trial >

Explore docs

sentiments.ts

write_sql.go

parse_expenses.py

addresses.rb

```
1 #!/usr/bin/env ts-node
2
3 import { fetch } from "fetch-h2";
4
5 // Determine whether the sentiment of text is positive
6 // Use a web service
7 async function isPositive(text: string): Promise<boolean> {
8   const response = await fetch(`http://text-processing.com/api/sentiment/`, {
9     method: "POST",
10    body: `text=${text}`,
11    headers: {
12      "Content-Type": "application/x-www-form-urlencoded",
13    },
14  });
15  const json = await response.json();
16  return json.label === "pos";
17 }
```

Copilot

Replay

HyperCLOVA (Korean Version of GPT-3) of NAVER

NAVER CLOUD PLATFORM

문의하기 콘솔

< AI Services | AITEMS CLOVA AiCall CLOVA Chatbot CLOVA Dubbing CLOVA OCR CLOVA Speech CLOVA Studio CLOVA Voice CLOVA S

CLOSED BETA

CLOVA Studio Update

CLOVA Studio는 초대규모 AI HyperCLOVA 기반의 No Code AI 도구입니다.

이용 신청하기

특징 상세기능 활용 예시 이용 안내 사용 가이드

CLOVA Studio를 활용해 다양한 문제를 해결하고, 누구나 쉽게 AI를 개발할 수 있습니다.

초대규모 AI HyperCLOVA
CLOVA Studio

캐릭터와 대화하기

데이터 생성하기

기획전 구성하기

음식 레시피 만들기

역사적 인물과 대화하기

콘텐츠 분류하기

마케팅 문구 생성하기

셀럽 말투 따라하기

비속어 필터하기

메뉴 추천해주기

상품 리뷰 요약하기

검색결과 요약

의도 이해하기

뉴스 요약해주기

반려동물 지식백과

엔티티 인식하기

소셜 쓰기

가사 쓰기

사투리 변환하기

백과사전 문고답하기

WRTN: Document Auto-Completion based on HyperCLOVA

The screenshot displays the WRTN web interface for creating an SNS advertisement. On the left is a sidebar with navigation options like '즐거찾기', '단계별 글쓰기', '블로그 포스팅', '책 초안', and '카피라이팅'. The main area is titled '제목 SNS 광고문구' and includes a search bar. The '제품 / 브랜드 이름' field contains '네이버 AI Lab', which is circled in green. Below it, the '제품 간단 정보' field contains '아시아 최고의 AI 연구소'. The '포함 키워드 (선택)' section shows selected keywords: '선행연구', '미래연구', '글로벌', '인공지능', and '아시아 최고'. At the bottom, there's a '생성 개수' dropdown set to '3' and a '자동 생성' button. On the right, the generated text is shown under the '전체' tab, with a '좋아요' button. A tooltip points to the '네이버 AI Lab' text, asking '결과가 마음에 드시나요?' and suggesting to click the '자동 생성' button again. The generated text includes: '구글과 페이스북보다 더 똑똑한 인공지능 개발중인 이 곳, 바로 국내 최초 초대 규모 AI전문가 조직 '네이버 AI LAB' 입니다.' and '당신에게 가장 좋은 것만 드리는'네이버 AI LAB에서는 선행연구와 미래연구 그리고 글로벌 인공지능과 아시아 최고 수준의 기술력을 갖춘 인재들과 함께 하고 있습니다.'

Powered by CLOVA 문의하기

제목 SNS 광고문구

제품 / 브랜드 이름 ? 전체 지우기

네이버 AI Lab

제품 간단 정보 ?

아시아 최고의 AI 연구소

포함 키워드 (선택) ? * 키워드를 한개 씩 작성 한 후 Enter를 눌러주세요.

코드

선행연구 × 미래연구 × 글로벌 × 인공지능 × 아시아 최고 ×

생성 개수 3 자동 생성

전체 좋아요 더 나은 결과물을 원하세요?

구글과 페이스북보다 더 똑똑한 인공지능 개발중인 이 곳, 바로 국내 최초 초대 규모 AI전문가 조직 '네이버 AI LAB' 입니다.

"당신에게 가장 좋은 것만 드리는"네이버 AI LAB에서는 선행연구와 미래연구 그리고 글로벌 인공지능과 아시아 최고 수준의 기술력을 갖춘 인재들과 함께 하고 있습니다.

세계적 수준의 인공지능 기술 개발 및 적용 사례가 궁금하다면? 지금 바로 NAVER AI LAB 홈페이지와 블로그를 방문해보세요.

Summary

- Models are getting bigger and bigger.
- Owing to self-supervised learning techniques, the language generation capability is getting better and better.
- We are getting closer to artificial **general** intelligence.