

Deep Learning Week 6

Language models: BERT & GPT

Hanyang Artificial Intelligence Group



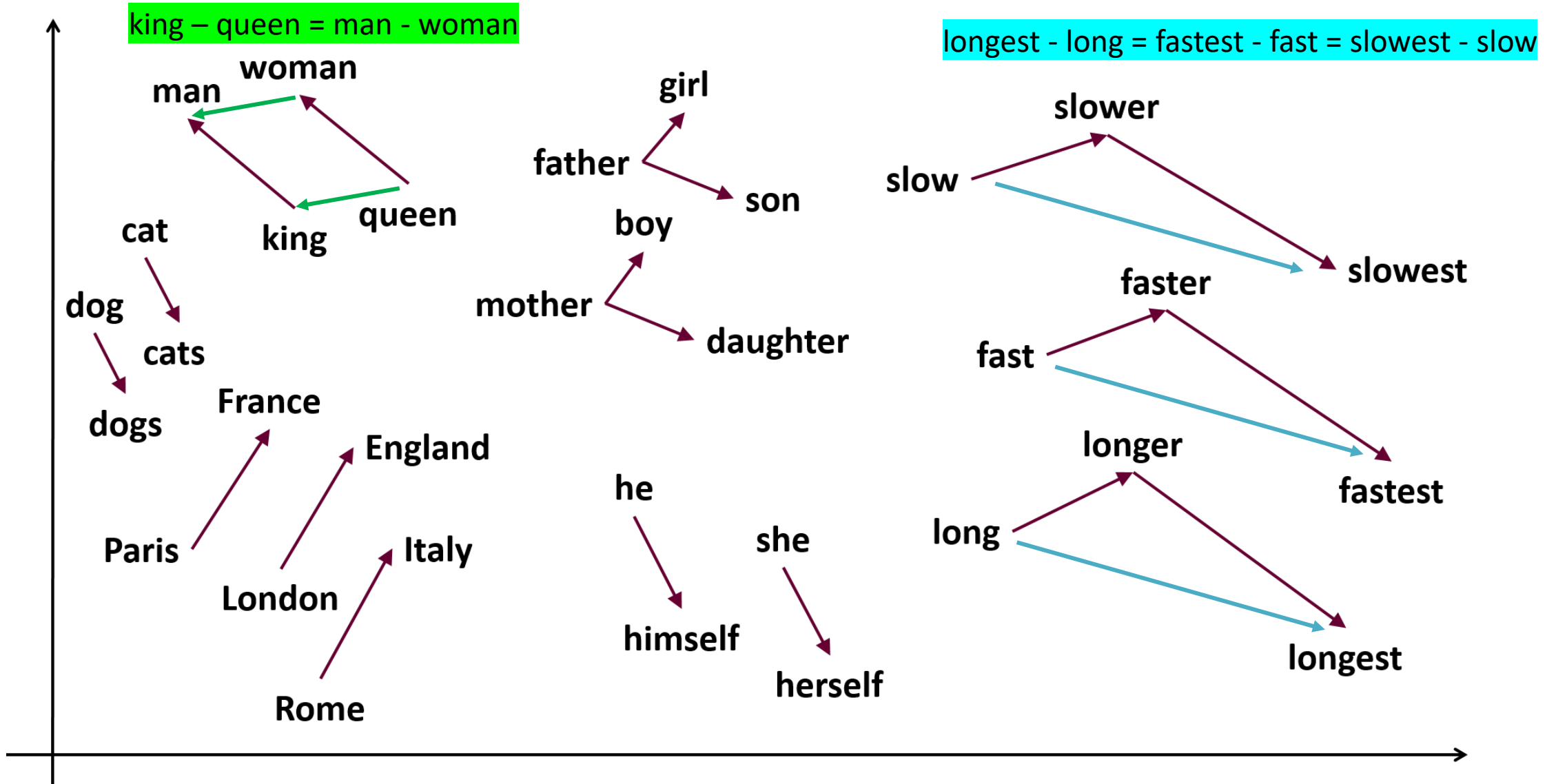
What is NLP?

인간의 언어를 이해할 수 있는 AI

- 인간이 일상에서 사용하는 언어의 의미를 분석하여 컴퓨터가 처리할 수 있도록 하는 기술
- 딥러닝을 활용한 방식의 발전으로 최근 크게 떠오르고 있는 분야
- NLP의 핵심 목표는 **자연어를 컴퓨터가 이해할 수 있는 표현**으로 변환하는 것!



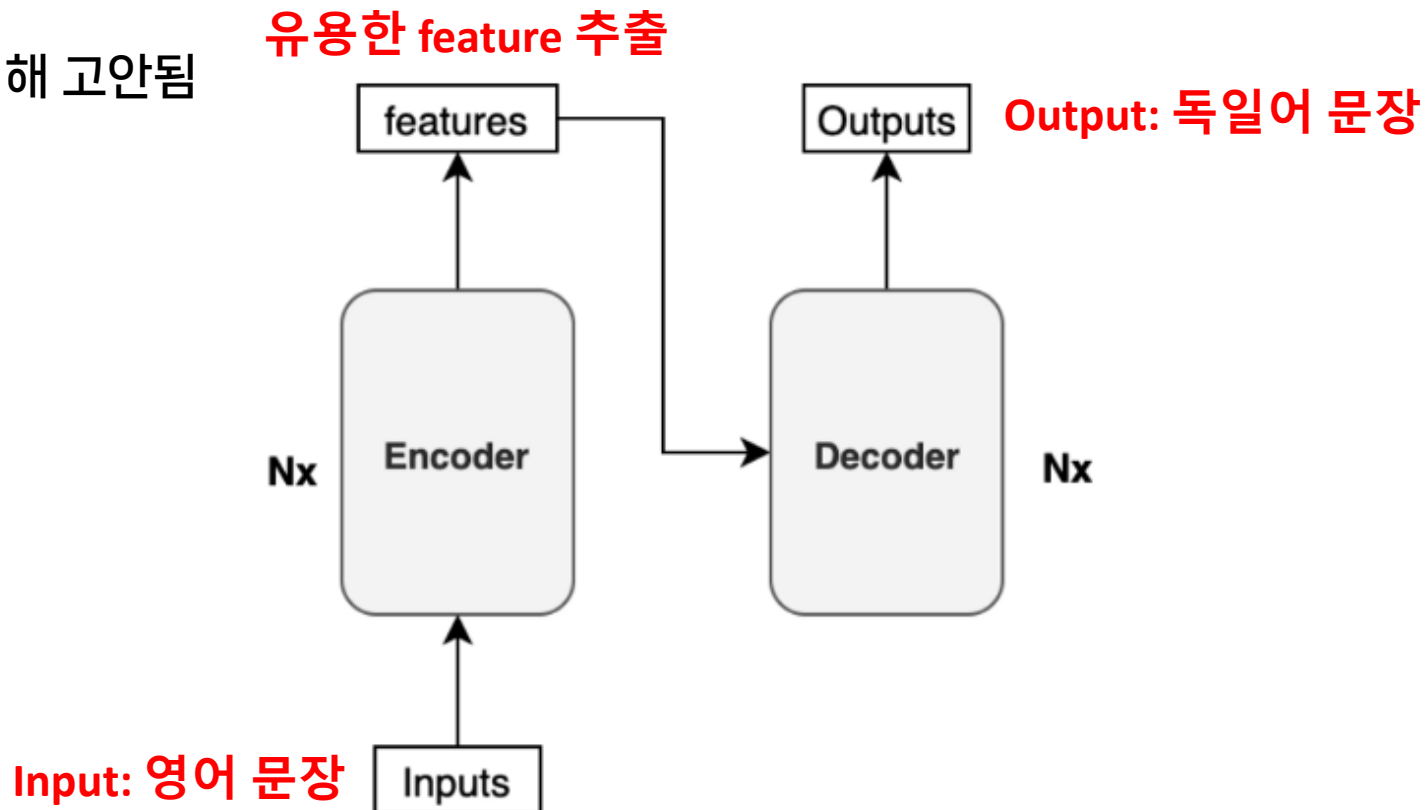
단어 임베딩 시각화 예시(Word2Vec)



Sequence-to-sequence

Seq2Seq: 입력을 문장으로 받아, 다시 문장을 출력할 수 있는 모델

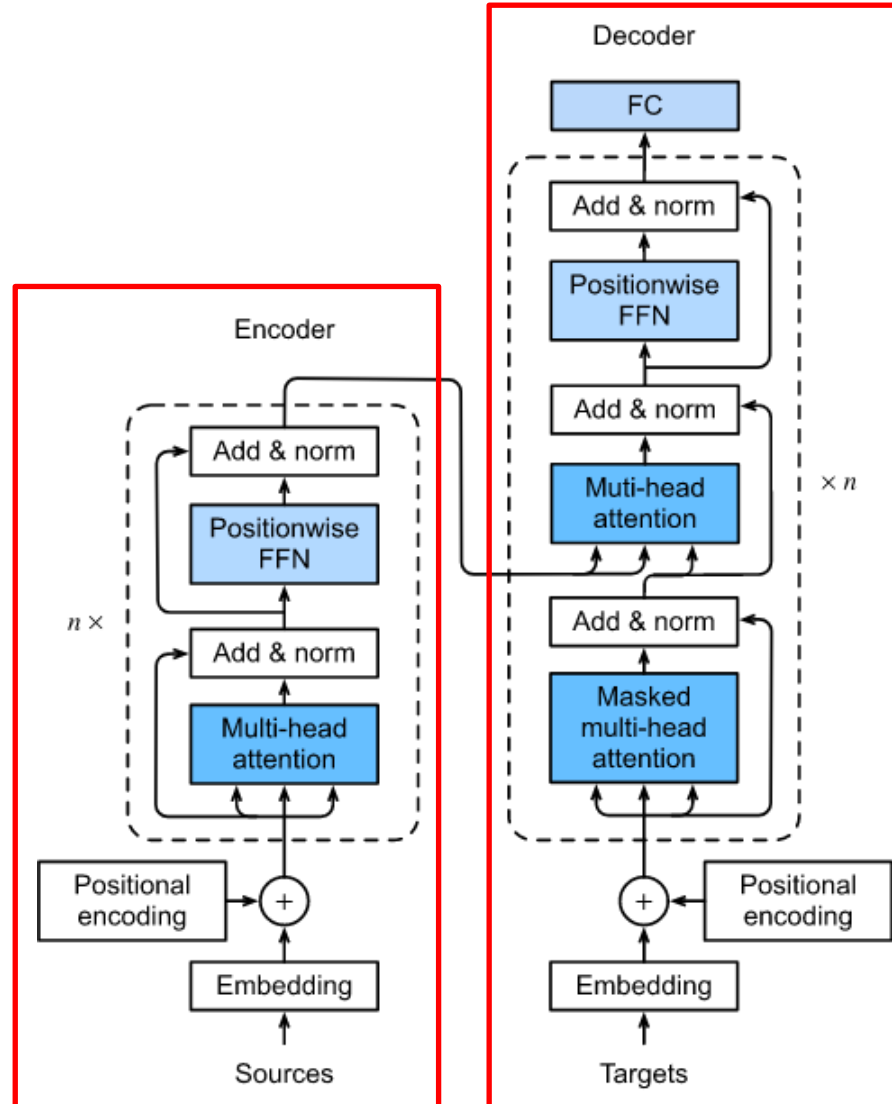
- 입력을 적절한 표현으로 변환하는 인코더(Encoder)와 변환된 정보를 바탕으로 새로운 데이터를 생성하는 디코더(Decoder)모델로 구성
- 기계 번역 등의 작업을 수행하기 위해 고안됨



BERT와 GPT



- Transformer의 Encoder만 사용
- 입력된 텍스트를 feature vector로 변환하는 능력
- **이해**에 뛰어난 모델



- Transformer의 Decoder만 사용
- 입력된 텍스트 다음에 이어질 내용을 예측하는 능력
- **생성**에 뛰어난 모델

What is a Language Model?

언어 모델: 주어진 문장에 대한 확률을 계산하는 함수!

- 언어 모델은 가장 자연스러운 단어 시퀀스를 찾아내는 모델로, 입력된 문장이 얼마나 자연스럽고 적절한지에 대한 확률을 계산하거나, 다음에 이어지는 단어를 예측하는 등의 작업을 수행할 수 있음



언어 모델 활용 예시: 검색 엔진

- 입력된 검색어 뒤에 이어지는 단어를 예측
- 각각의 단어에 대한 확률을 기준으로 가장 가능성이 높은 순서대로 정렬해서 보여줌

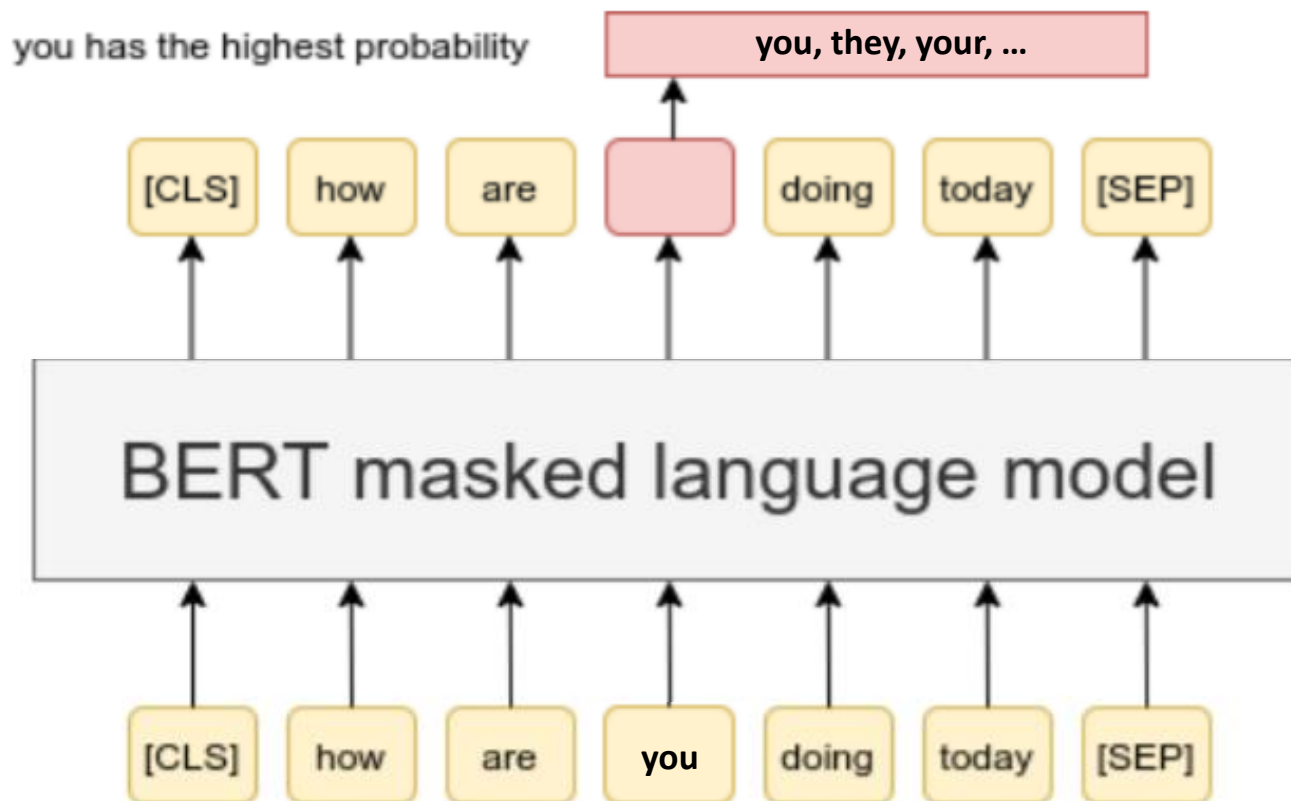
딥 러닝을 이용한 |

딥 러닝을 이용한 부동산가격지수 예측
딥 러닝을 이용한 자연어 처리 입문
딥 러닝을 이용한 한국어 의존 구문 분석
딥 러닝을 이용한 개체명 인식
딥 러닝을 이용한 차량 번호판 검출
딥 러닝을 이용한 한국어 의미역 결정
딥 러닝을 이용한 한국어 형태소의 원형 복원 오류 수정
딥 러닝을 이용한
딥 러닝을 이용한 구문 분석

BERT의 Pretraining 방법: Masked Language Modeling

- 원본 텍스트의 일부 단어를 가리고 해당 위치에 가장 적절한 단어를 맞추도록 학습하는 과정
- 데이터에 사람이 labeling을 진행할 필요 없이, 모델이 텍스트만 가지고도 스스로 언어의 특성을 학습할 수 있게 됨

-> Self-supervised Learning



- 다음 빈칸에 들어갈 말로 가장 적절한 것을 고르시오.

Since human beings are at once both similar and different, they should be treated equally because of both. Such a view, which grounds equality not in human uniformity but in the interplay of uniformity and difference, builds difference into the very concept of equality, breaks the traditional equation of equality with similarity, and is immune to monist distortion. Once the basis of equality changes so does its content. Equality involves equal freedom or opportunity to be different, and treating human beings equally requires us to take into account both their similarities and differences. When the latter are not relevant, equality entails uniform or identical treatment; when they are, it requires differential treatment. Equal rights do not mean identical rights, for individuals with different cultural backgrounds and needs might [redacted] in respect of whatever happens to be the content of their rights. Equality involves not just rejection of irrelevant differences as is commonly argued, but also full recognition of legitimate and relevant ones. [3점]

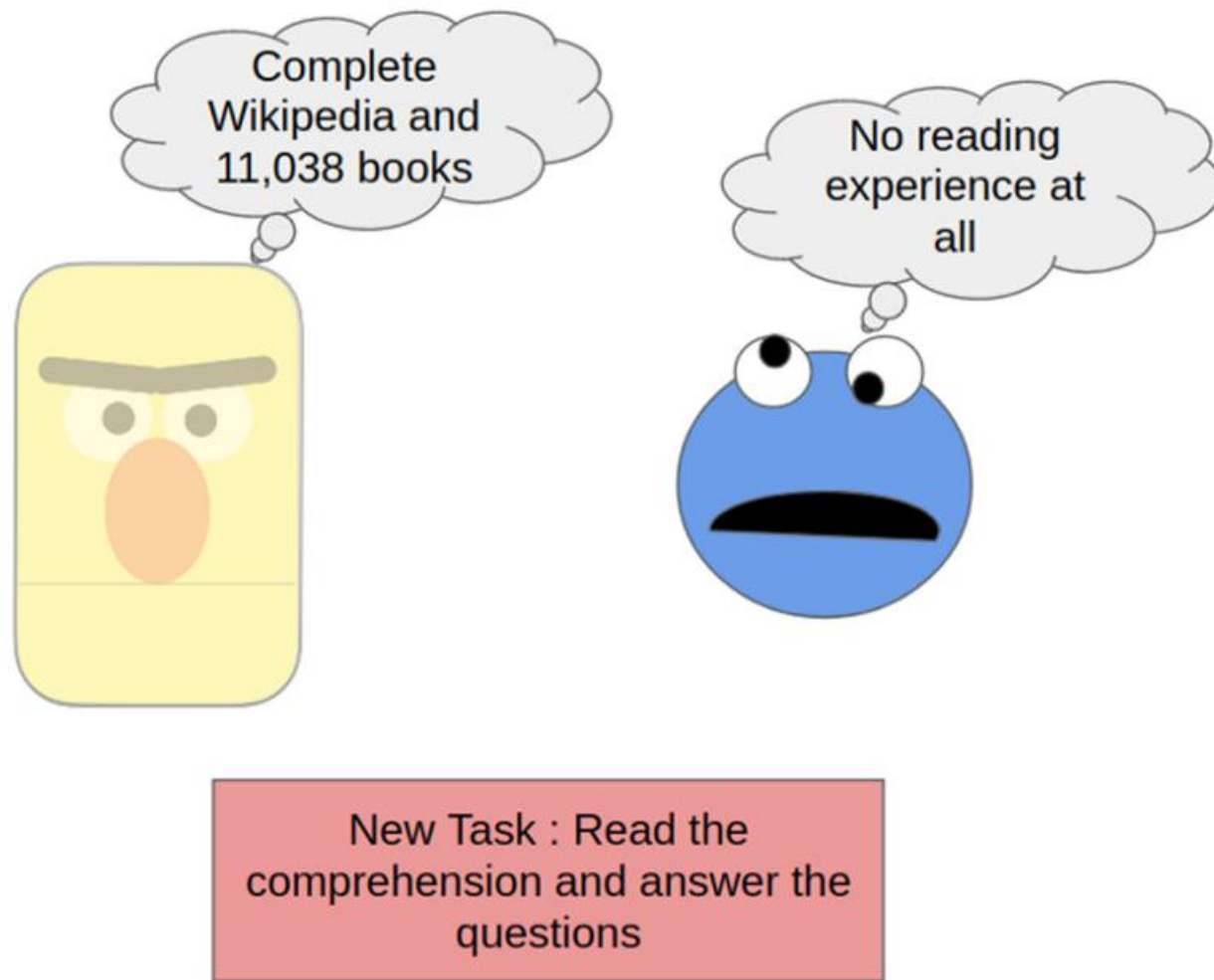
* monist: 일원론의 ** entail: 내포하다

- ① require different rights to enjoy equality
- ② abandon their own freedom for equality
- ③ welcome the identical perception of inequality
- ④ accept their place in the social structure more easily
- ⑤ reject relevant differences to gain full understanding

Recall: Transfer learning

이미 학습된 모델의 능력을 재사용한다!

- 유사한 task에 대해 학습된 적 있는 모델의 능력을 활용하여 새로운 task를 해결하도록 적용하는 방법
- **Pretraining:** 모델이 입력 데이터의 특성을 잘 학습할 수 있도록 학습하는 과정
- **Finetuning:** 사전 학습된 모델을 원하는 다른 task를 수행하도록 학습하는 과정



BERT의 Transfer Learning 과정

Text corpus



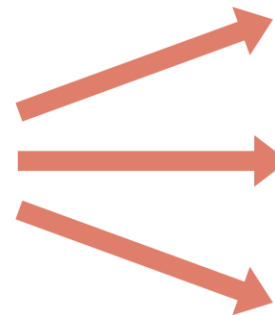
(Self-supervised)
Training



Pretrained LM



Adaptation



Tasks

Question
Answering



Text
Classification



Information
Retrieval

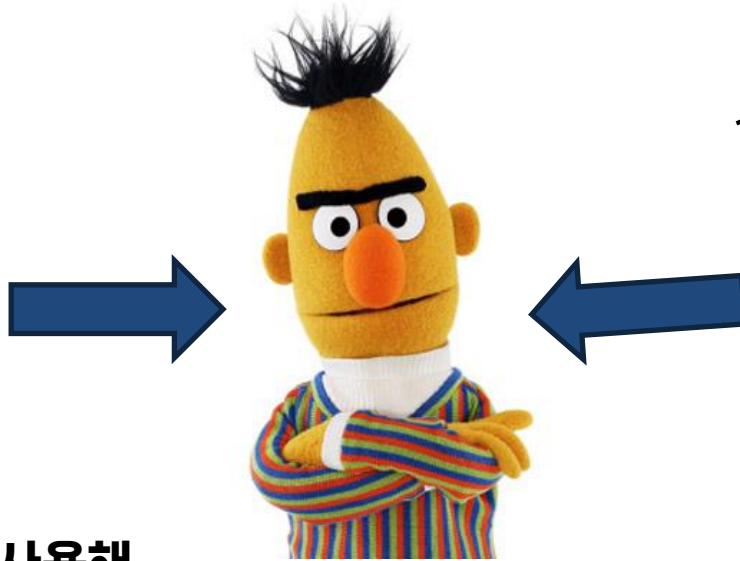


⋮

BERT를 활용한 Text Classification



Step 1. 한국어 데이터를 사용해
MLM pretraining 진행



ㅋㅋㅋㅋ 진짜 너무 웃겨 ㅋㅋ **긍정**
완전 막장임... 이것도 영화라고 만드냐 **부정**
완전 막장임... 돈날렸어 ㅋㅋ **부정**
안 봤으면 큰일날 뻔. 주인공이 너무 예쁘다 **긍정**
완전 너무 재밌다!! 인생영화! **긍정**
개노잼... 다시는 이 감독 영화 안 봐야지 **부정**

Step 2. 긍정/부정 라벨링된 영화 리뷰
데이터로 finetuning 진행

꼭 한 번 더 보고 싶어요!

긍정

영화 리뷰 감정 분석 모델 완성!

BERT를 활용한 Sentence Embedding

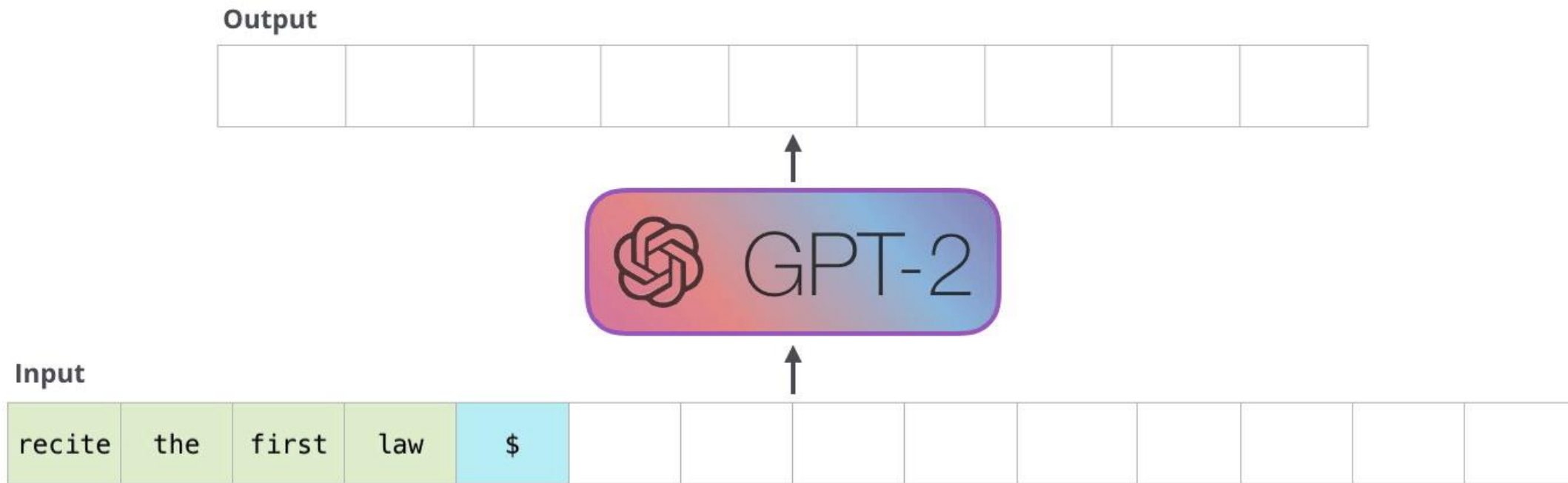


Sentence Embedding

비슷한 의미를 가진 문장들끼리 가깝게 배치되도록
벡터 형태로 변환하는 과정

GPT의 Pretraining 방법: Next Token Prediction

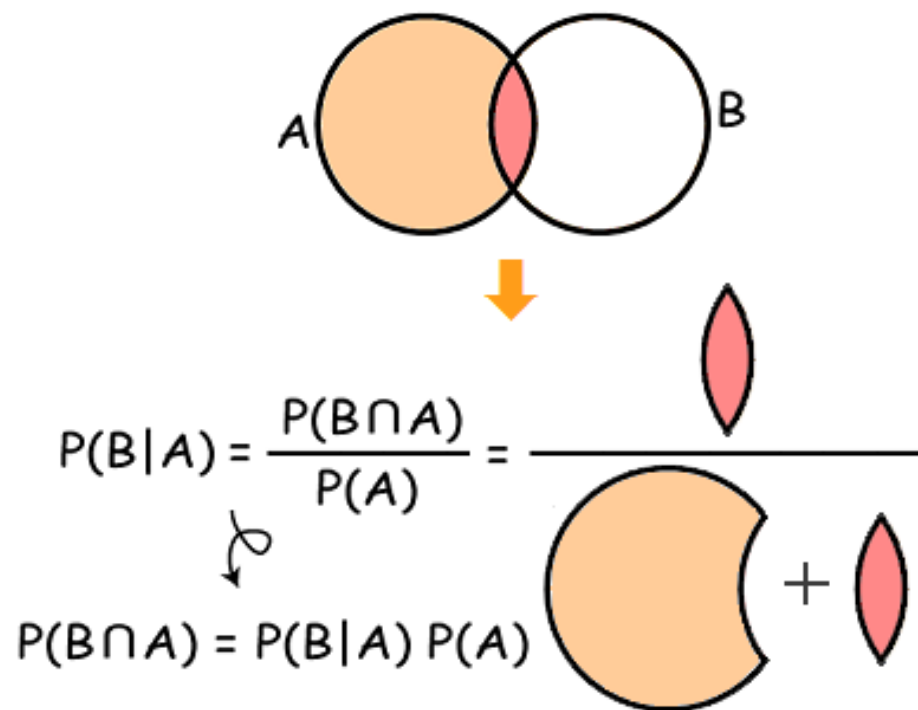
- 원본 텍스트에 이어지는 가장 적절한 단어를 맞추도록 학습하는 과정
 - 데이터에 사람이 labeling을 진행할 필요 없이, 모델이 텍스트만 가지고도 스스로 언어의 특성을 학습할 수 있게 됨
- > **Self-supervised Learning**



조건부 확률과 GPT의 Language Modeling 방법

GPT 모델은 조건부 확률을 계산하는 함수!

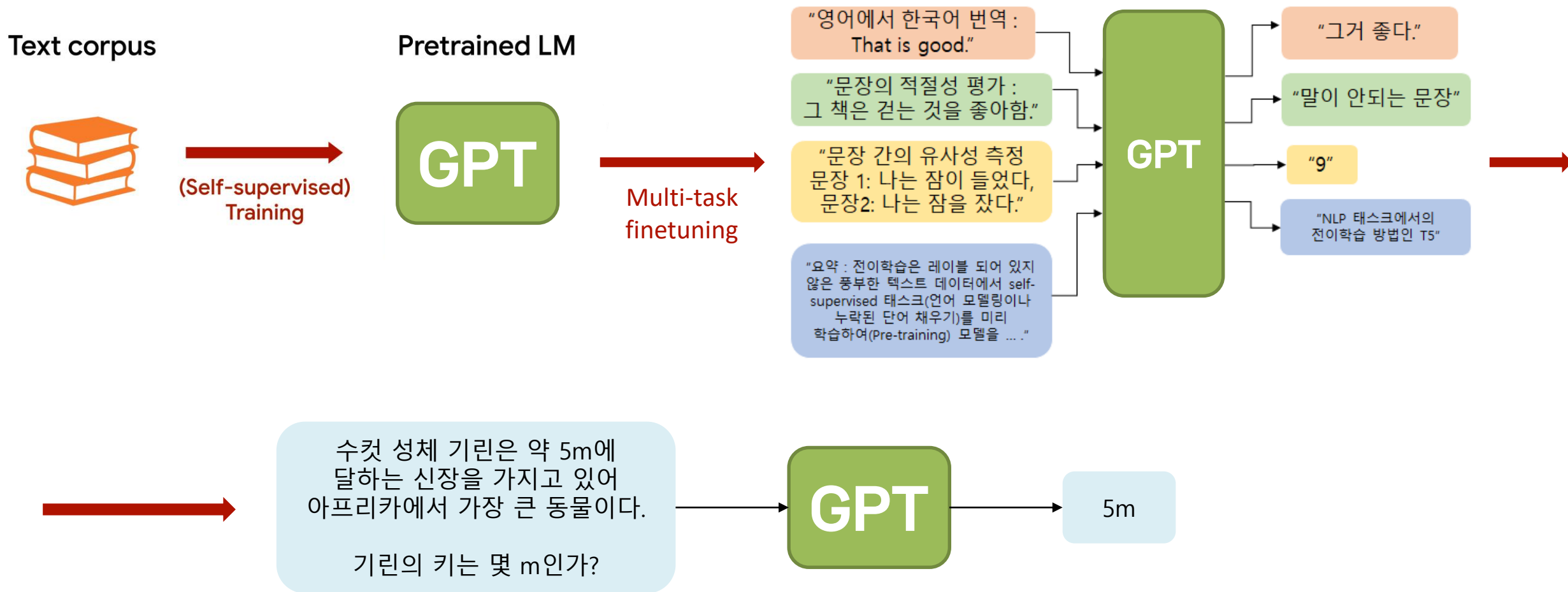
- Next token prediction 과정에서 언어 모델은 입력된 문장에 이어지는 단어 후보군마다 각각의 확률을 부여하게 됨
- 입력된 문장에 이어지는 글을 생성하는 것은 후보 단어들의 확률을 기반으로 선택하는 과정!



GPT("오늘 내가 먹은 ")
= P(X | "오늘 내가 먹은")

$$= \begin{cases} 0.122 & X = \text{"음식"} \\ 0.094 & X = \text{"메뉴"} \\ 0.084 & X = \text{"과자"} \\ 0.042 & X = \text{"식단"} \\ 0.012 & X = \text{"것"} \\ \dots & \end{cases}$$

GPT의 Transfer Learning 과정



학습한 적 없는 task에 대해서도 좋은 성능을 보여준다!

Prompt Engineering이란?

원하는 결과물을 얻기 위해선 프롬프트에 집중하라!

- 프롬프트 엔지니어링은 질문을 잘 하여 AI를 효과적으로 쓰는 방법
- AI로부터 최상의 답변을 끌어내기 위하여 최적의 단어, 또는 입력 값(프롬프트)의 조합을 찾는 작업

Standard Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The answer is 27. ❌

Chain-of-Thought Prompting

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✅

프롬프트를 이용한 학습: In-Context Learning

Finetuning 없이 원하는 task를 가르쳐보자!

- **Few shot learning:** 프롬프트에 원하는 task를 수행하는 방법에 대한 예시를 함께 주면, **추가적인 finetuning 없이** 모델이 task를 해결하는 능력을 가지게 되는 것

LS

"사자"로 이행시를 지어줘.



무한한 대지에 휘날리는 털의 묵직함
눈부신 갈기로 세상을 지배하며
눈빛은 태양을 비추고
포식자에게는 위협의 군주로 다가섭니다.

무리 속에는 용맹한 사자들이 함께 춤을 춥니다.
함성은 하늘에 울려 퍼지며
사자의 힘은 대지를 흔들어 놓쳐버릴 정도로 강력합니다.
주목하라, 사자의 왕이여!
자부심과 자유로움이 가득한 사자의 시대가 지금 시작됩니다.

LS

"하마"로 이행시를 지어줘.

하: 하마는 입이 참 크다.
마: 마음은 그만큼 넓지 않은 것 같다.

"표범"으로 이행시를 지어줘.
표: 표범은 참 멋있는 무늬가 있다.
범: 범상치 않은 외모가 인상적이다.

"사자"로 이행시를 지어줘.



사: 사자는 대자연의 왕이다.
자: 자신감과 위엄이 넘치는 모습이다.

GPT의 문제점: Hallucination(환각)

GPT에게 사실 기반의 판단을 맡기기 위험한 이유

- GPT 모델은 텍스트를 생성할 때 데이터에서 학습한 확률 기반으로 다음 단어를 예측하는 과정만 반복하기 때문에, 생성된 텍스트가 사실과는 전혀 다른 내용일 수 있음

조선왕조실록에 기록 된 세종대왕의 맥북프로 던짐 사건에 대해 알려줘

세종대왕의 맥북프로 던짐 사건은 역사서적인 조선왕조실록에 기록된 일화로, 15세기 조선시대 세종대왕이 새로 개발한 훈민정음(한글)의 초고를 작성하던 중, 문서 작성 중단에 대한 담당자에게 분노하여 맥북프로와 함께 그를 방으로 던진 사건입니다.

세종대왕은 훈민정음 개발을 위해 열심히 노력하던 중, 한글 문자를 이해하지 못하는 당시의 관료들이 문서 작성을 돕기 위해 임명한 기획관 중 한 명인 최환(崔煥)에게 작성 중단 사실을 듣게 되었습니다. 세종대왕은 이에 분노하여 최환을 부르고, 맥북프로를 함께 들고 그를 방으로 불러들이며 화를 내었습니다. 이후, 세종대왕은 훈민정음의 초고 작성을 다른 담당자에게 넘겨주게 되었습니다.

여섯 번째 과제: 🤖 Transformers로 BERT와 GPT 사용하기

Step 1. 예시 노트북을 실행해보며 BERT와 GPT 모델 사용해보기

- 🤖 transformers 라이브러리를 사용하여 공개된 모델을 불러와 사용하는 방법에 대해 확인해 봅시다.
- BERT는 텍스트를 이해하는 task(분류 문제 등), GPT는 텍스트를 생성하는 task(번역, 요약 등)에 뛰어납니다.
- Task에 적합한 모델을 Huggingface hub에서 찾아 다운로드해서 사용해 봅시다.

Step 2. Sentence BERT 모델을 사용하여 문장 사이의 유사도 계산하기

- Sentence BERT는 문장 임베딩(텍스트 -> 벡터로 변환)을 수행할 수 있는 모델입니다.
- Sentence BERT는 유사한 의미를 가지는 문장의 임베딩 벡터 사이의 거리가 가깝도록 학습되었습니다.
- Cosine similarity, Euclidean distance 등을 활용하여 문장과 문장 사이의 유사도를 계산해 봅시다.
- 참고 자료(Weekly NLP): <https://jiho-ml.com/weekly-nlp-5/>

With HAI, Fly High

Hanyang Artificial
Intelligence