

LLM

Large Language Model은

P(text)를 계산 할 수 있는 큰 모델

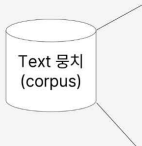
텍스트의 확률 계산 모델
Language Model

일반적으로는 Transformer

Language Model

- 원래의 Language Modeling은 Unsupervised learning / Self-supervised learning으로 텍스트의 확률을 계산 할 수 있도록 하는 방법
- 대량의 Raw Text Corpus에 대해 Language Modeling을 학습하는 것으로 Pretraining 도 가능
- Language Model을 텍스트의 확률을 계산하는 본연의 의미보다 Representation learning의 용도로 사용하는 트렌드가 있었음

Language Modeling



고양이는 귀여운 동물입니다
강아지도 귀엽습니다
AI가 과연 사람의 일자리를 없앨까요?
답러닝은 AI의 한 분야입니다.
...

Language Modeling - Text 문치에 있는 모든 텍스트의 출현 확률을 학습

P(text)

Factorization of joint probability

- 예) [1] '고양이'는 단어뒤에 출현할 단어의 확률중 **가장 큰 확률**을 가진 '귀여운'을 대답
- 예) [2] '고양이는 귀여운' 문장뒤에 출현할 단어의 확률중 **가장 큰 확률**을 가진 '동물입니다'를 대답

$$P(\text{고양이는 귀여운 동물입니다}) = \begin{matrix} P(\text{고양이는}) \\ \times \\ 1 \quad P(\text{귀여운} \mid \text{고양이는}) \\ \times \\ 2 \quad P(\text{동물입니다} \mid \text{고양이는 귀여운}) \end{matrix}$$

GPT vs Gemini

<https://arxiv.org/pdf/2312.11805.pdf>

모델	GPT	Gemini
개발자	OpenAI	Google
모델 아키텍처	Transformer Decoder +alpha	Transformer Decoder +alpha
입력 모달리티	텍스트, 이미지	텍스트, 오디오, 이미지, 비디오(이미지의 연속)
출력 모달리티	텍스트	텍스트, 이미지
성능	다양한 벤치마크에서 인간 전문가 능가	다양한 벤치마크에서 인간 전문가 능가
모델 크기 및 유형	gpt-3.5-turbo, gpt-4-turbo, gpt-4	Ultra, Pro, Nano
개발자 지원 및 접근성	API를 통한 접근, GPTs	Android 개발자를 위한 Gemini Nano 프리뷰, Gemini Pro API 접근성

대략 알파벳 4자 당 1token 한국어는 1자당 1~2토큰

모델	GPT-3.5-turbo	GPT-4-turbo	GPT-4	Gemini pro
입력	\$0.001 / 1,000 토큰 approx. \$0.00025 / 1,000자(영문)	\$0.01 / 1,000 토큰 approx. \$0.0025 / 1,000자(영문)	\$0.03 / 1,000 토큰 approx. \$0.0075 / 1,000자(영문)	\$0.00025 / 1,000자(영문 기준)
출력	\$0.002 / 1000 토큰 approx. \$0.0005 / 1,000자(영문)	\$0.03 / 1000 토큰 approx. \$0.0075 / 1,000자(영문)	\$0.06 / 1,000 토큰 approx. \$0.015 / 1,000자(영문)	\$0.0005 / 1,000자(영문)

<https://platform.openai.com/tokenizer>

프롬프트

이 사진을 바탕으로 좋고 흥미로운 블로그 게시물을 작성해 주세요. 사진에 음식에 대한 설명이 포함되어야 하고, 음식 준비에 관해 이야기해야 합니다.



응답

식사 준비는 시간과 비용을 절약할 수 있는 좋은 방법일 뿐만 아니라, 더 건강한 음식을 먹는 데 도움이 될 수 있습니다. 이 식사는 미리 쉽게 준비할 수 있는 건강하고 맛있는 식사의 좋은 예입니다.

편의, 구운 야채, 달고기 데리야케가 포함된 식사입니다. 특이한 건강하고 복잡한 한우를로, 지속적인 에너지를 공급합니다. 구운 채소는 비타민과 미네랄을 매일 섭취하는 데 좋은 방법이며 저탄 데리야케는 맛있고 단백질이 풍부한 예입니다.

이 식사는 준비도 매우 쉽습니다. 냄비와 팬을 구우고 저탄 데리야케를 요리하기만 하면 됩니다. 모든 것이 조리되면 식사 준비 용기로 나누어 냉장고에 보관합니다. 그런 다음 컨테이너를 들고 비싼 아침을 줄일 수 있습니다.

이러한 쉽게 준비할 수 있는 건강하고 맛있는 식사를 찾고 있다면 이 식사가 좋은 선택입니다. 영양가가 가득하고 맛이 있기 때문에 포만감과 만족감을 느낄 수 있습니다.

여기에서는 건강하고 맛있는 식사를 준비할 수 있습니다!

→ 프롬프트 설계 기초

- 구체적으로 안내: 오해의 여지가 최소화되도록 명확하고 간결한 안내를 작성합니다.
- 프롬프트에 몇 가지 예시 추가: 현실적인 퓨트 예시를 사용해 달성하려는 목표를 설명하세요.
- 단계별 세분화: 복잡한 작업을 관리 가능한 하위 목표로 나누어 프로세스 전반에 걸쳐 모델을 안내합니다.
- 출력 형식 지정: 프롬프트에서 마크다운, JSON, HTML 등 원하는 형식으로 출력하도록 요청합니다.
- 단일 이미지 프롬프트를 위해 이미지를 먼저 배치: Gemini는 이미지와 텍스트 입력을 원하는 순서로 처리할 수 있지만 단일 이미지가 포함된 프롬프트의 경우 해당 이미지를 텍스트 프롬프트 앞에 배치하면 성능이 더 좋을 수 있습니다.[]

구체적으로 안내의 예

프롬프트



이 이미지를 설명해주세요.

모델응답

공항 도착 및 출발 안내판을 보여주는 이미지입니다.

업데이트된 프롬프트



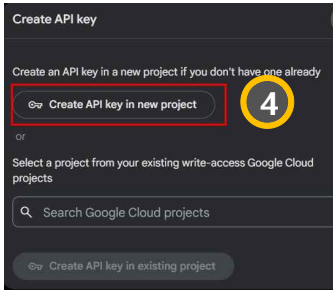
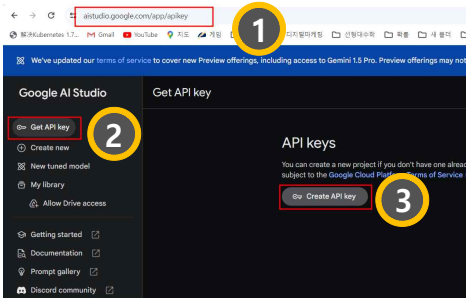
이 이미지에 표시된 공항 보드의 시간과 도시를 목록으로 파싱합니다.

개선된 대응

- 10:50 모스크바
- 11:05 예든버러
- 11:05 런던
- 11:10 부카레스트
- 11:30 키예프
- 11:35 더블린
- 11:45 이스터 밀렌즈
- 12:15 소피아
- 12:30 워던
- 12:30 뉴캐슬
- 12:40 상트페테르부르크
- 12:40 런던
- 12:45 맨체스터

구글 바드에서 API키를 발급받아서 사용합니다.

https://aistudio.google.com/app/prompts/new_chat



기본폴더- 1기본_gemini_api.ipynb

```
1 # Or use os.getenv('GOOGLE_API_KEY') to fetch an en
2 GOOGLE_API_KEY="<YOUR GOOGLE API KEY>"
3 genai.configure(api_key=GOOGLE_API_KEY)
```