

State of GPT

20220301

No.

Andrej Karpathy

고흥기

GPT 훈련과정

1. pre-training ① 계산 집약적, 데이터 수집
훈련 시간의 99% 차지.

② Tokenization

text \rightarrow GPT가 작동하는 기본 표현인 토큰 sequence로
byte encoding

③ 프로, train

1750억 개의 매개변수

④ 작동방식

트랜스포머에 token batch 공급 \rightarrow 다음 token 예측 시도
Loss function으로 예측 정확도 측정.

2. Supervised ① 작고 고품질 data set 수만개 생성

Fine tuning ② Prompt - 응답 쌍 생성

(SFT) ③ 알고리즘: 사전 훈련과 동일하게 언어 모델링 작업 수행, 훈련 세트로 교체

④ 결과: 어느 정도 실제 어시스턴트로 작동 가능.

3. Reward ① data: 비교 대상으로 data 수집

Modeling ② train: transformer가 각 응답의 품질에 대한 보상 예측하도록 훈련.

4. 강화학습 ① 작동방식: 보상 모델 사용, Prompt에 대한 응답에 대해

(Reinforcement Learning) 점수 매김. \rightarrow 모델 매개변수 조정

② 결과: ChatGPT가 RLHF 모델 예시

from RLHF > SFT.

Human Feedback ③ 단점: RLHF는 엔트로피가 높음

RLHF) \hookrightarrow 피크가 높은 결과 제공

\hookrightarrow 더 낮은 변이 출력

\hookrightarrow 기본 model보다 다양성 \downarrow

LLM 활용방법

No.

1. LLM 특징 ① 단순히 다음 token 모방

② 수백억 개 매개변수, 거대한 작업 memory

2. Prompt ① Chain-of-Thought

engineering : 추론의 분산, 단계별로 생각.

② Self-consistency

: 한번만 시도하는데 아님, 여러번 sampling 해서 여러번 시도

③ self Reflection

: 자신이 실수를 저질렀는지 스스로 확인

④ Tree of Thought

: 잘 되는 Prompt 는 계속 유지

⑤ RAG (Retrieval Augmented Generation)

: 관련 문서를 참고로 하고, 벡터 임베딩하고, 벡터 db에 저장하고,

쿼리할때 관련 참고 검색

Fine Tuning ① 세기 : Prompt engineering 을 훨씬 잘하고 생각할 때.

고려사항 ② 장점 : LoRa 와 같이 모델의 작은 부분 훈련하면

효율성 높일 수 있다.

③ 단점 : 기술적으로 훨씬 더 작업이 필요.

특히 RLHF는 연구 영역이 매우 넓고 범위가

적용하기가 매우 어렵다.

LLM의 한계 ① bias 편향 : 데이터에 따른 편향

② Hallucination & Fabrication (조작)

③ Reasoning error (추론 오류)

④ knowledge cut-off : 특정 날짜 이후의 data X

⑤ 보안 취약

권장사항 → 인간 감독 필수, LLM이 자율적인 agent 가 아닌 보조자

MOCKEUK
Capitol