

대규모 언어 모델의 활용이 지속적으로 발전하면서 발생할 수 있는 윤리적인 고려사항

제목 : Language Models are Few-Shot Learners

논문 내용

GPT-3라는 대규모 언어 모델의 성능과 가능성에 대해 말함.

2.GPT-3 소개

GPT-3는 1750억 개의 파라미터로 구성된 자기회귀형 언어 모델로, 이전의 언어 모델보다 훨씬 많은 파라미터를 가지고 있음. 이 모델은 미세 조정 없이도 다양한 자연어 처리(NLP) 작업에서 강력한 성능을 보이며, 기존의 최첨단 미세 조정 방법과 경쟁력을 갖추고 있다함.

3.Few-Shot 성능 향상

논문은 GPT-3가 과제에 종립적인(few-shot) 성능을 크게 향상시킨다는 것을 보여줌. 이 모델은 적은 수의 예시만으로도 성능을 테스트할 수 있으며, 텍스트 상호작용을 통해 다양한 NLP 작업을 수행할 수 있음.

4.다양한 NLP 작업에서의 성능

GPT-3는 번역, 질문 응답, 문장 내 새로운 단어 사용, 3자리 산술 등 다양한 NLP 데이터셋에서 강력한 성능을 보임.

또한, 논문은 사람들이 구분하기 어려운 뉴스 기사도 생성할 수 있음을 보여주고 있음.