

1. GPT

GPT 모델은 일론 머스크와 샘 알트만이 설립한 openAI 에서 개발한 자연어 처리 모델입니다. GPT 가 특별한 이유는 언어모델을 사용하기 때문인데, 이를 사용하면서 라벨링이 필요가 없어졌습니다. 사람이 직접 라벨링을 할 경우 비싸고 시간도 오래걸리고 정확도도 떨어지기에 이는 큰 장점입니다. 하지만 GPT 모델이 완벽한 것은 아닙니다. GPT-2 는 단방향(unidirectional) 언어 모델로서, 입력 프롬프트 다음에 나오는 단어를 예측하도록 훈련이 되어 있습니다. 이는 양방향(bidirectional) 언어 모델인 BERT 에 비해 단점이 뚜렷하다는 지적을 받았습니다. 하지만, 2019 년에 단방향 언어 모델인 GPT-2 의 진가를 보여주는 작업이 있었는데, 바로 텍스트 생성 작업이었습니다. 인공지능을 이용한 텍스트 생성 작업에 있어서 단방향 모델이 양방향 모델보다 뛰어난 성능을 보여준 것 입니다. GPT-1 은 1 억 1 천 7 백만개의 파라미터를 가지고 있고, GPT-2 는 GPT-1 보다 약 12.8 배 많은 15 억 4 천 2 백만개의 파라미터를 가지고 있고, 약 800 만개의 문서, 40GB 용량의 데이터셋을 학습시킨 모델입니다.

2. BERT

트랜스포머의 인코더를 쌓아올린 구조입니다. Pre-trained 된 bert 위에 classification layer 를 하나 추가해주면 다양한 nlp 를 처리할 수 있다는 장점이 있습니다. BERT 는 2019 년 10 월 25 일 구글 리서치 팀에 의해 공개된 자연어처리 사전 훈련 모델입니다. BERT 모델은 100 여개가 넘는 언어 학습을 지원하며, BERT-Base, BERT-Large, BERT-Base, Multilingual, 그리고 BERT-Base, Chinese 모델이 있습니다. 각각의 모델 뒤에 Cased 와 Uncased 가 붙혀져 있는데, Uncased 의 경우 대소문자 구분을 하지 않는 모델입니다. BERT 를 이용하여 특정 과제를 수행할 수 있는데요, 이를 위해서는 세부적인 과제를 수행하도록 파인튜닝(fine-tuning) 작업이 필요합니다. BERT 은 사전 학습된 언어모델로서 파인튜닝을 통해, 원하는 작업을 수행할 수 있습니다. BERT 가 발표 되었을 당시, GPT-2 와 비교를 많이 했는데요, 두 모델 모두 파인튜닝(fine-tuning) 방식을 사용하지만, GPT-2 는 단방향(unidirectional) 언어 모델인 반면, BERT 는 양방향(bidirectional) 언어 모델로서 차이가 있었습니다. 양방향 언어 모델의 장점은 fill-in-the-blanks 와 같이 앞뒤 문맥에 맞게 빈칸에 알맞은 단어를 추측하는 작업에 아주 높은 성능을 보여주고 있습니다.