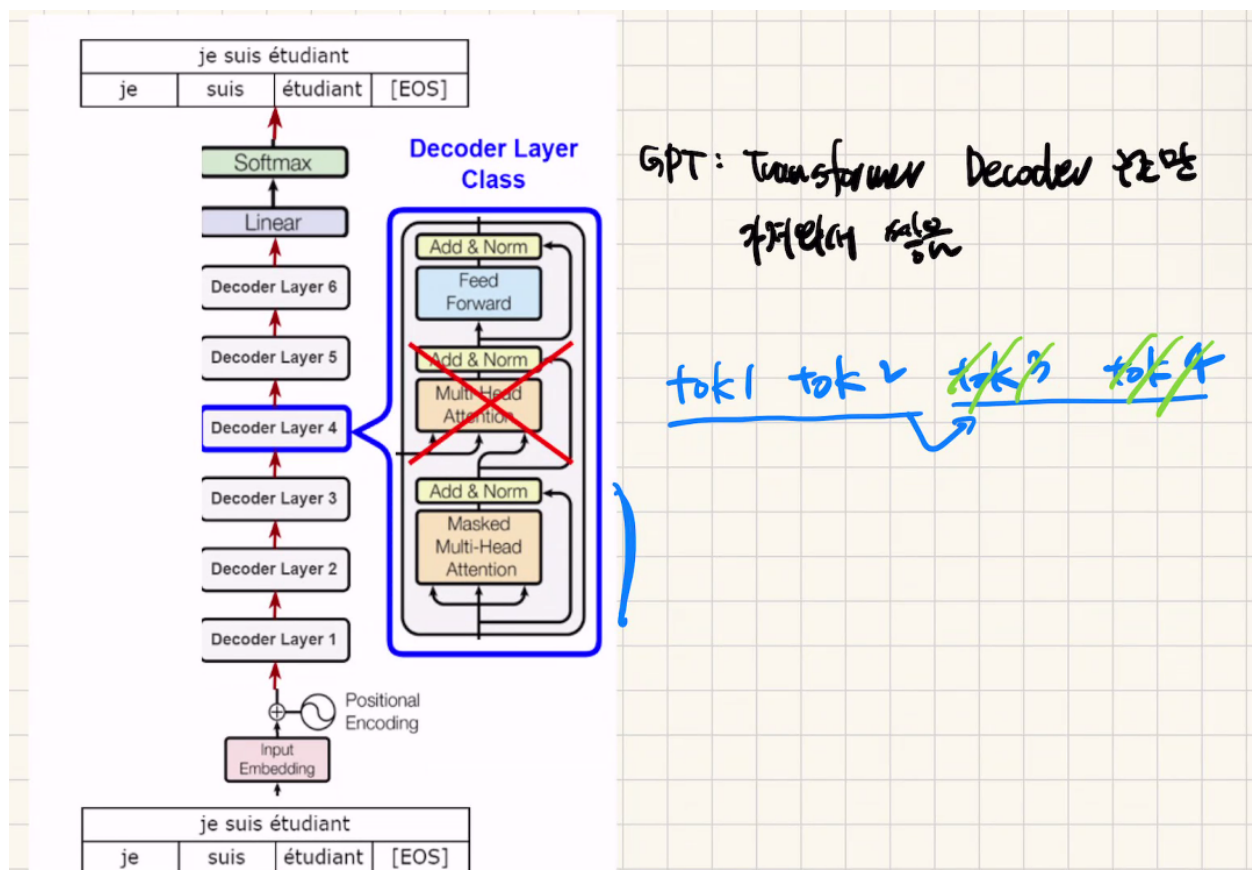


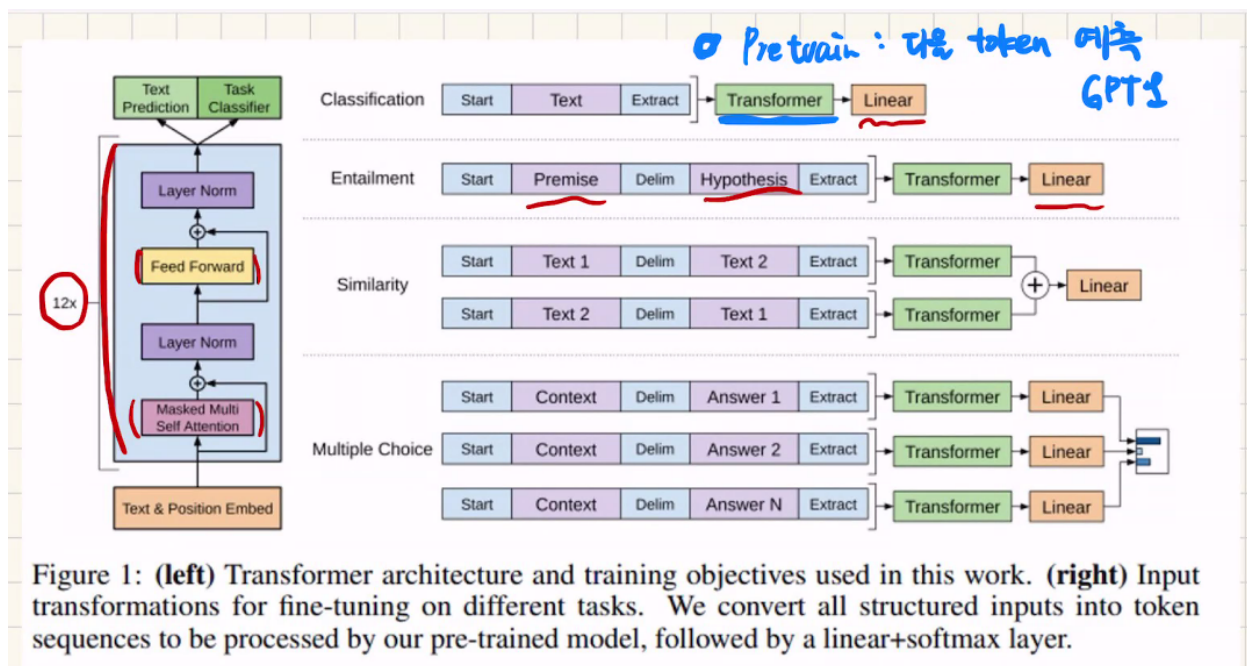
GPT (다음단어맞추기가 핵심)

practice 최신트랜드 논문 · 2024. 10. 7. 14:38

Transformer의 decoder부분만 씀

pretrain: 다음 token 예측 -> GPT1





♡ 2



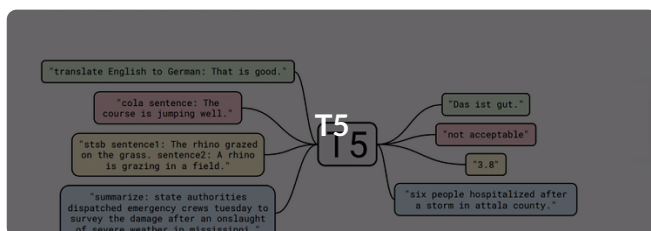
구독하기

'practice 최신트랜드 논문' 카테고리의 다른 글

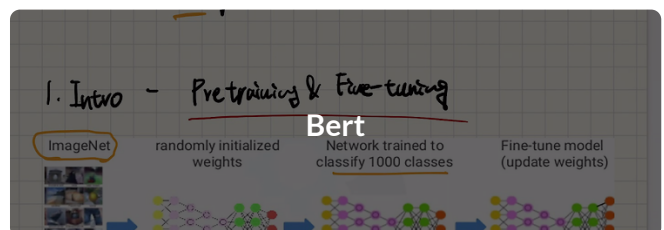
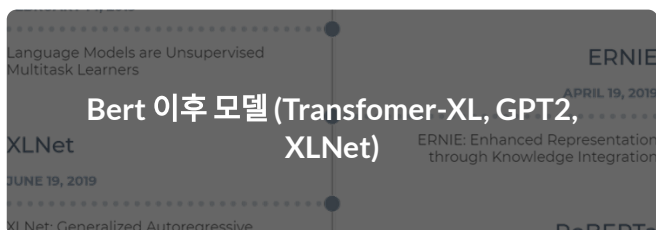
T5 (1)	2024.10.12
RoBERTa (0)	2024.10.12
Bert 이후 모델 (Transfomer-XL, GPT2, XLNet). (2)	2024.10.12
Bert (0)	2024.10.04

관련글

관련글 더보기



Model	data	bsz	steps	SQuAD (v1.1/2.0)	MNLI-m	SST-2
RoBERTa						
with BOOKS + WIKI	16GB	8K	100K	93.6/87.3	89.0	95.3
+ additional data (§3.2)	160GB	8K	300K	94.0/87.7	89.3	95.6
+ pretrain longer	160GB	8K	500K	94.4/88.7	90.0	96.1
+ pretrain even longer	160GB	8K	500K	94.6/89.4	90.2	96.4
BERT _{LARGE}						
with BOOKS + WIKI	13GB	256	1M	90.9/81.8	86.6	93.7
XLNet _{LARGE}						
with BOOKS + WIKI	13GB	256	1M	91.0/82.8	88.4	94.4



자연어(NLP)

네이쳐2024 님의 블로그입니다.



구독하기 +

댓글 1



mwollossna



공감~!남기고 갑니다^^

2024. 10. 7. 22:53 · [답글](#)



이름

비밀번호

내용을 입력하세요.



댓글 남기기