주제 : 타이거 모델 경량화

V100 fine-tuning 1~2시간

실험 환경 : GCP 인스턴스, docker

기한 : 2주

필요 작업 :

\*base 모델 분석

\*base 모델 경량화 비교군 성능 조사(경량 transformer, 경량 bert)

\*base 모델 블록화

\*경량/pruning기술 조사

\*Base 모델/블록 학습 및 그래프 확보

\*Base 모델/블록 성능 측정

블록 별 inference time, 전체 정확도

(Base 블록 output feature 데이터화)

(Output feature 데이터 분석)

\*경량 모델/블록 학습 및 그래프 확보

\*경량 모델/블록 성능 측정

블록 별 inference time, 정확도,

\*결과 정리