**자연어처리 실습 과제 2 보고서**

**1. 요약 :**

sshleifer/distilbart-xsum-12-3 모델에서 facebook/bart-large로 모델을 변경하여 동일한 실험를 진행, loss 값을 4.3223에서 3.7031까지 낮출 수 있었음

**2. 실험 진행 :**

**모델 변경 -** facebook/bart-large 모델 사용

텍스트, 폰트, 스크린샷이(가) 표시된 사진

자동 생성된 설명

**수행 Task -** 텍스트 요약(Summarization)

**사용 데이터셋 –** wiki\_lingua

**평가 함수** - ROUGE 점수 계산 함수

**3. 결과 :**

sshleifer/distilbart-xsum-12-3 모델

**텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명**

'eval\_loss': 4.212963104248047

Facebook/bart-large 모델

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

'eval\_loss': 3.703291350415015

**3-1 결과 해석 :**

**3-1-1 "sshleifer/distilbart-xsum-12-3"와 "facebook/bart-large" 모델의 차이점 :**

**모델 구조 및 크기 – “**sshleifer/distilbart-xsum-12-3”모델은 원래의 BART 모델을 경량화한 버전. XSum 데이터셋을 위해 12개의 인코더 레이어와 3개의 디코더 레이어로 구성되어 있음.원래 BART보다 더 적은 매개변수(계산량이 적음)를 사용하여 더 빠른 속도로 학습하고 예측.

“facebook/bart-large” 모델은 원래의 BART Large 모델로, 12개의 인코더 레이어와 12개의 디코더 레이어로 구성되어 있음. 더 큰 모델 구조와 매개변수를 통해 복잡한 문장 구조를 학습하는 능력이 뛰어남.

**훈련 데이터 및 사전 학습 –** “sshleifer/distilbart-xsum-12-3” 모델은 XSum 데이터셋을 요약 작업용으로 미리 학습한 버전.

“facebook/bart-large” 모델은 더 일반적인 대규모 코퍼스를 사용해 미리 학습된 사전 학습 모델. 다양한 NLP 작업에 일반화된 능력을 가지며, 다양한 데이터셋과 작업에서 높은 성능을 발휘할 수 있음.

**3-1-2 비교 결과 :**

요약 Task에서도 "facebook/bart-large" 모델이 "sshleifer/distilbart-xsum-12-3" 모델보다 더 큰 구조와 사전 학습 데이터를 기반으로 하여 향상된 성능을 보여줌.

.

**4. 한계 :**

다양한 데이터셋과 모델에 대해서도 실험해보고 싶었으나 계산량이 커질 경우 Colab의 GPU사용량 초과로 연결이 끊기는 현상이 빈번히 발생하게 되어 가장 직관적으로 평가가 가능한 실험 방식을 채택함.