

**SK네트웍스 Family AI과정 3기  
 모델링 및 평가 테스트 계획 및 결과 보고서**



**□ 개요**

* 산출물 단계 : 모델링 및 평가
* 평가 산출물 : 테스트 계획 및 결과 보고서
* 제출 일자 : 2024.12.24 (화)
* 깃허브 경로 : https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN03-FINAL-1Team
* 작성 팀원 : 진윤화

| **개요** | * LLM 텍스트 임베딩 모델 성능 평가를 위한 MTEP 벤치마크 테스트 결과 보고 * 선행 논문 추천 기능을 구현하기 위해, 원 논문과 선행 논문의 요약을 text embedding을 통해 벡터로 구현하고, cosine similarity를 구함. |
| --- | --- |
| **MTEB 벤치마크** | * text embedding에 사용될 모델의 성능을 평가하기 위해 MTEB의 STS16 task를 활용함 * MTEB 벤치마크를 통해 snowflake-arctic-embed-l-v2.0, all-mpnet-base-v2, nomic-embed-text-v1.5 모델들을 테스트함. |
| **MTEB 벤치마크 결과** | * 각 모델별 main score는 다음과 같음 * snowflake-arctic-embed-l-v2.0: 0.834902 * all-mpnet-base-v2: 0.800299606531299 * nomic-embed-text-v1.5: 0.834902 * 위의 결과를 통해 snowflake-arctic-embed-l-v2.0, nomic-embed-text-v1.5 모델을 선택함. |
| **실제 텍스트 적용** | * snowflake-arctic-embed-l-v2.0, nomic-embed-text-v1.5 모델을 실제 텍스트에 적용한 결과, nomic-embed-text-v1.5모델은 한국어 모델에 적절한 성능을 보이지 못하여 snowflake-arctic-embed-l-v2.0을 최종 선택함. |
| **결과** | * MTEB 벤치마크를 통해 snowflake-arctic-embed-l-v2.0, nomic-embed-text-v1.5 모델을 선택하였으나, 한국어 지원 이슈로 인하여 snowflake-arctic-embed-l-v2.0을 최종 선택함. |