

프로젝트 주제	경제 지식이 없는 국민들을 대상으로 경제 지식을 쉽게 설명하는 AI 인플루언서 웹 페이지 관리 Tool
사용 데이터	<b>[ 한국은행 제공 경제교육용 PDF 문서 ]</b> - 경제금융용어 700선 - 알기 쉬운 경제 이야기 (일반인용)
인공지능 학습 목표	<p>해당 프로젝트에서는 경제 지식이 없는 이들을 위해 어려운 용어를 쉽게 풀어서 작성하는 것을 목표로 합니다.</p> <p>따라서 '변환 텍스트(Output)'는 '원본 텍스트(Input)'의 의미를 해치지 않는 선에서 쉬운 어휘로 패러프레이징한 형태여야 합니다.</p> <p>인공지능 학습 결과를 측정하기 위해 사용하는 정량 지표는 '코사인 유사도 (Cosine Similarity)'입니다.</p> <p>코사인 유사도는 -1 이상 1 이하의 값을 가지며, 1에 가까울수록 동일한 내용의 텍스트에 근접한다고 판단할 수 있습니다.</p> <p>그러나 변환 텍스트와 원본 텍스트를 비교했을 시에 1 값을 갖게 되는 것이 무조건 긍정적인 것은 아닙니다.</p> <p>텍스트 간의 전후 차이가 없을 때 1 값이 나오기 때문입니다.</p> <p>따라서 0.7 ~ 0.8 사이의 코사인 유사도 값이 나올 경우를 긍정적인 케이스라 판단하기로 했습니다.</p> <p>측정 샘플로는 랜덤으로 고른 30개 데이터를 사용했습니다.</p> <p>이는 경제금융용어 700선, 알기 쉬운 경제 이야기 데이터 모두 동일합니다.</p> <p><b>[ 요약 ]</b></p> <ul style="list-style-type: none"> <li>- 해당 프로젝트에서는 0.7 ~ 0.8 사이의 코사인 유사도를 긍정적으로 생각</li> <li>- 측정 샘플로 30개를 사용해서 각각의 코사인 유사도 값을 구한 후, 평균값으로 최종 산출</li> </ul>
인공지능 학습 결과 비교 과정	<p>인공지능 학습 결과를 정량적으로 측정하기 위해서 모델 구조를 최대한 단순화시킨 형태로 baseline 코드를 작성했습니다.</p> <p>원래 목표로 하던 LangGraph 구조는 LangChain 구조에 측정 지표와 if ~ else 구문 그리고 node 등을 첨가하여 순환 구조를 만든 형태이기 때문에, LangChain 기본 성능을 먼저 따진 후에 추후 고도화 작업에 임하기로 결정했습니다.</p> <p>인공지능 학습 결과 비교 과정은 아래와 같습니다.</p> <p><b>[ 인공지능 학습 결과 비교 과정 ]</b></p> <ol style="list-style-type: none"> <li>1. 말투 파인튜닝 학습 후 모델 저장</li> <li>2. 프롬프트 엔지니어링</li> <li>3. LangChain 구조 정의</li> <li>4. 원본 텍스트(Input)와 변환 텍스트(Output)에 각각 동일한 임베딩 방식 적용</li> <li>5. 임베딩한 텍스트 사이의 Cosin Similarity 구하기</li> <li>6. 샘플 30개 Cosin Similarity 값 사이에서 평균 구하기</li> </ol>
'경제금융용어 700선' 학습 모델 성능	<pre>np.mean(similarity_li)</pre> <p>0.759737979690981</p> <p>해당 모델은 0.7~0.8 사이의 결과값이 나왔으므로 프로젝트에서 추구하는 방향으로 학습되었다는 것을 확인할 수 있습니다.</p>
'알기쉬운 경제 이야기 (일반인용)' 학습 모델 성능	<pre>np.mean(similarity_li)</pre> <p>0.7656560878633116</p> <p>해당 모델은 0.7~0.8 사이의 결과값이 나왔으므로 프로젝트에서 추구하는 방향으로 학습되었다는 것을 확인할 수 있습니다.</p>
추후 학습 방향	위 과정을 베이스로 측정 지표, if ~ else 구문, node 등을 첨가하여 LangGraph 사용 방향으로 프로젝트를 고도화할 예정입니다.