

스포츠 관용 표현의 언어학적 특징과 AI 번역기의 의미 보존력 분석

AI를 활용한 영어 연구

20232501 백은주, 20232562 박은우, 20250856 김유민



목차

01 프로젝트 제작 의도

- 연구 배경 및 목적

02 언어학적 분석

- 스포츠 기사 vs 관용 표현의 품사 및 단어 길이 비교 분석 단계
- 결과 설명
- 연구 의의

03 번역 정확도 비교: DeepL vs ChatGPT

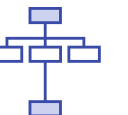
- 비교 실험 방법:
(1) 파일 업로드 방식
(2) 웹사이트 실험
- 결과 설명
- 번역기별 장단점 및 연구 의의



프로젝트 제작 의도



- 스포츠 관용 표현의 언어학적 특징 & 정확성 분석
- > 의미 왜곡의 우려가 있는 표현들을 올바르게 이해하는 데 도움을 주고자
이 프로젝트를 기획



언어학적 분석



- 스포츠 기사 vs 관용 표현의 품사 및 단어 길이 비교 단계



STEP 1
데이터 준비
및 환경 설정

텍스트 파일 업로드 및 확인
한글 폰트 설정 및 적용 확인
NLTK 리소스 다운로드



STEP 2
텍스트 전처리 및
분석 함수 정의

텍스트 정제 및 토큰화
품사 태깅 등 분석용 핵심 함수 구현



STEP 3
개별 텍스트
분석 수행

관용 표현 및 스포츠 기사
텍스트 분석 실행



STEP 4
분석 결과 시각화

품사 분포 그래프
단어 길이 분포 그래프



언어학적 분석



STEP 1. 데이터 준비 및 환경 설정

- 텍스트 파일 업로드 및 확인
 - files.upload() 사용하여 파일 업로드
- 한글 폰트 설정 및 적용 확인
 - Matplotlib을 사용하여 그래프를 그릴 때 한글이 깨지지 않고 올바르게 표시
- NLTK 리소스 다운로드
 - 영어 텍스트 분석에 필요한 NLTK(Natural Language Toolkit)의 주요 데이터 패키지들을 준비



언어학적 분석

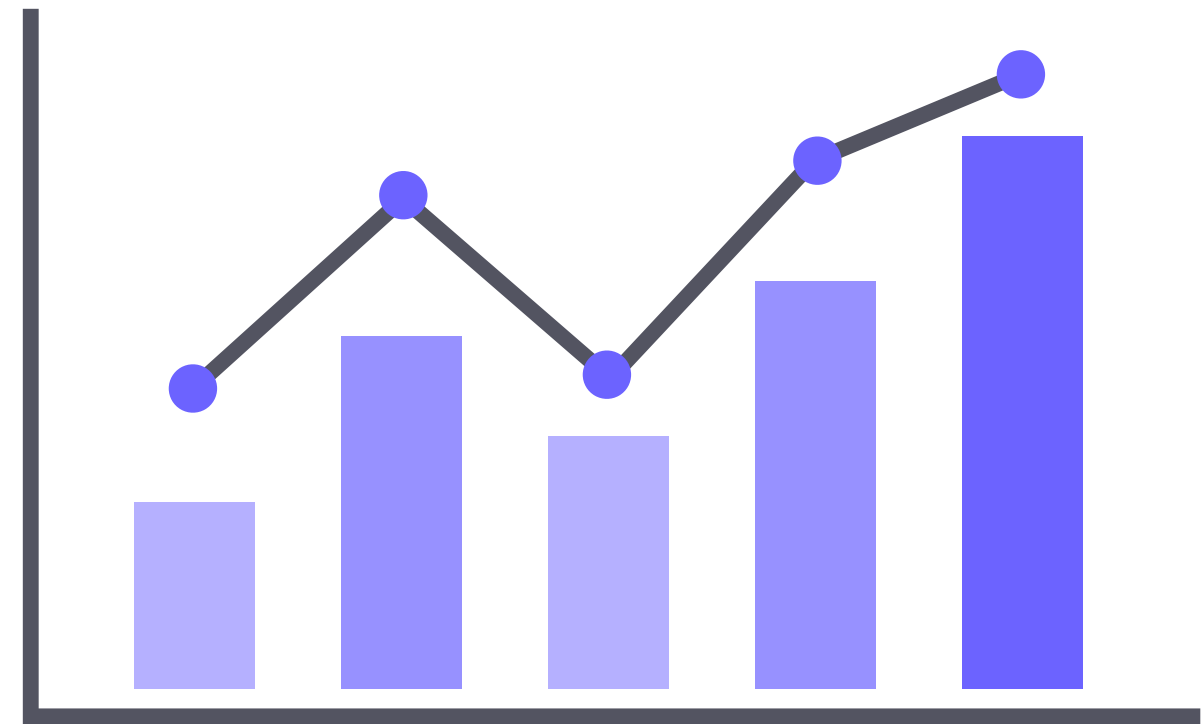


NLTK란? (Natural Language Toolkit)

-> 파이썬에서 텍스트를 처리하고 분석하기 위해 사용하는 도구 모음

- 주요 기능

- 토큰화
- 품사 태깅
- 불용어 제거



언어학적 분석



STEP 2. 텍스트 전처리 및 분석 함수 정의

- 텍스트 정제 및 토큰화
 - 텍스트 전처리 (preprocess_text): 불필요한 문자 제거, 토큰화, 불용어 처리, 표제어 추출 등을 통해 텍스트를 정제



STEP 3. 개별 텍스트 분석 수행

- 관용 표현 및 스포츠 기사 텍스트 분석 실행
 - analyze_pos (품사 분석), analyze_word_lengths (단어 길이 분석), analyze_noun_collocations (명사 연어 분석) 함수를 사용하여 개별 분석 결과를 도출



언어학적 분석



STEP 4. 분석 결과 시각화

- 품사 분포 그래프
 - 데이터 준비: 'Noun', 'Verb' 등 주요 카테고리로 단순화하고 빈도수를 집계
 - 시각화: 집계된 단순화 품사 카테고리별 빈도수를 막대그래프로 나타냄 (x축: 품사 카테고리, y축: 출현 빈도)
- 단어 길이 분포 그래프
 - 데이터 준비: 단어 길이별 빈도수를 길이 기준으로 정렬
 - 시각화: 길이 기준으로 정렬된 단어 길이별 빈도수를 막대그래프로 나타냄 (x축: 단어 길이, y축: 해당 길이의 단어 빈도수)



언어학적 분석



결과 설명

- 이용한 스포츠 기사

Bavarian Football Works. (2024, September 8). Germany faces tough Nations League draw with Hungary

<https://www.bavarianfootballworks.com/2024/9/8/24238491/germany-uefa-nations-league-hungary-gundogan-jamal-musiala-florian-wirtz-nagelsmann-club-world-cup>

- 이용한 관용표현

Pass the baton

Backed into a corner

Cross the finish line

Gain an early advantage

Increase the out count

Enter the ring

Hit a home run

Make a final spurt

Emerge as a dark horse

Make a gamble



언어학적 분석



idioms 품사 분포 결과

=====

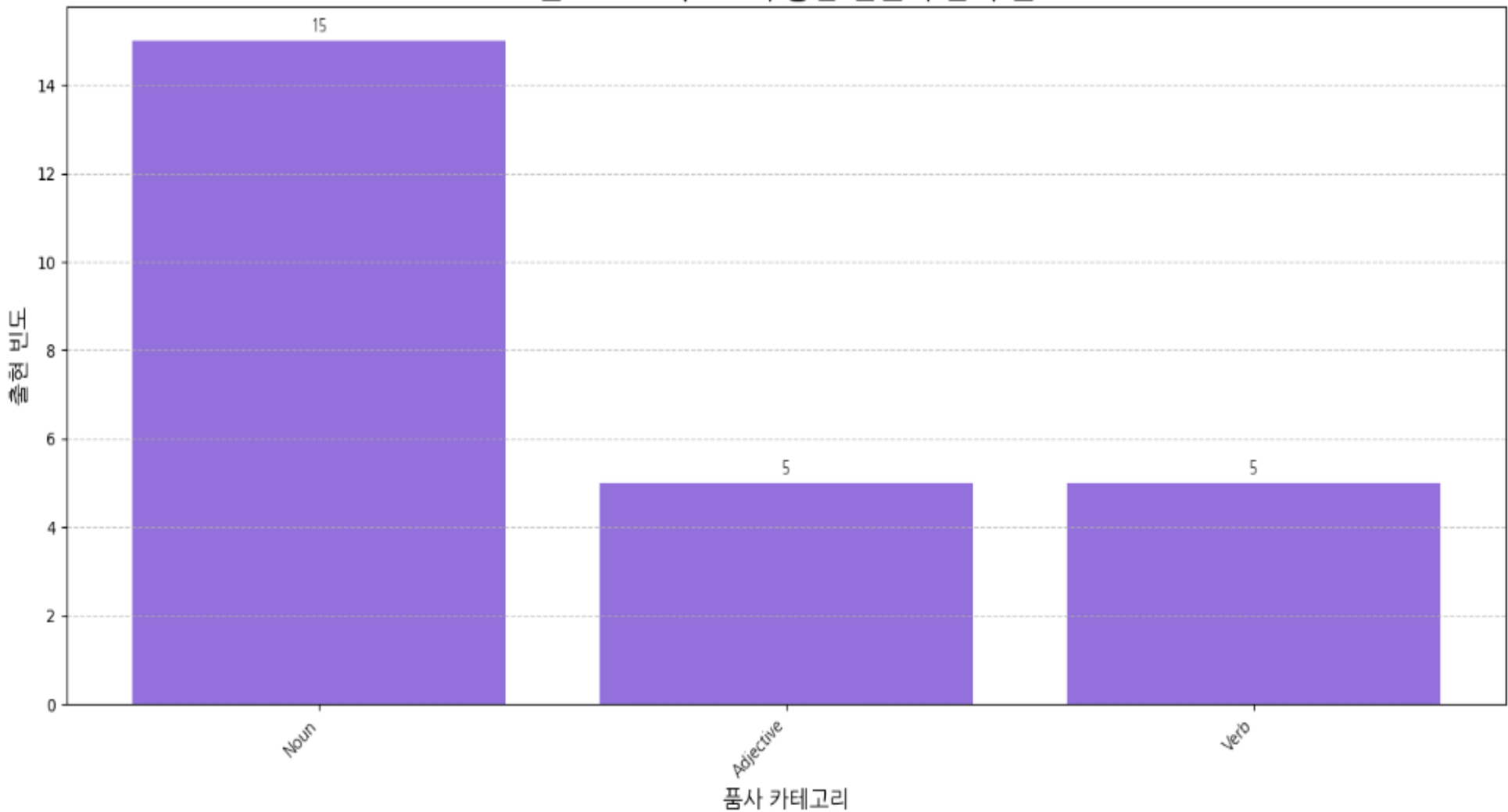
모든 Idiom 파일의 통합 품사 분포 시각화

=====

모든 Idiom 텍스트의 통합된 단순화 품사 카테고리 분포:

Noun: 15
Adjective: 5
Verb: 5

모든 Idiom 텍스트의 통합 단순화 품사 분포



- 가장 많이 등장하는 품사
- 명사(Noun) → 형용사(Adjective) = 동사(Verb)
- 주로 명사, 형용사, 동사가 관용구의 핵심 구성 요소



언어학적 분석



article의 품사 분포 결과

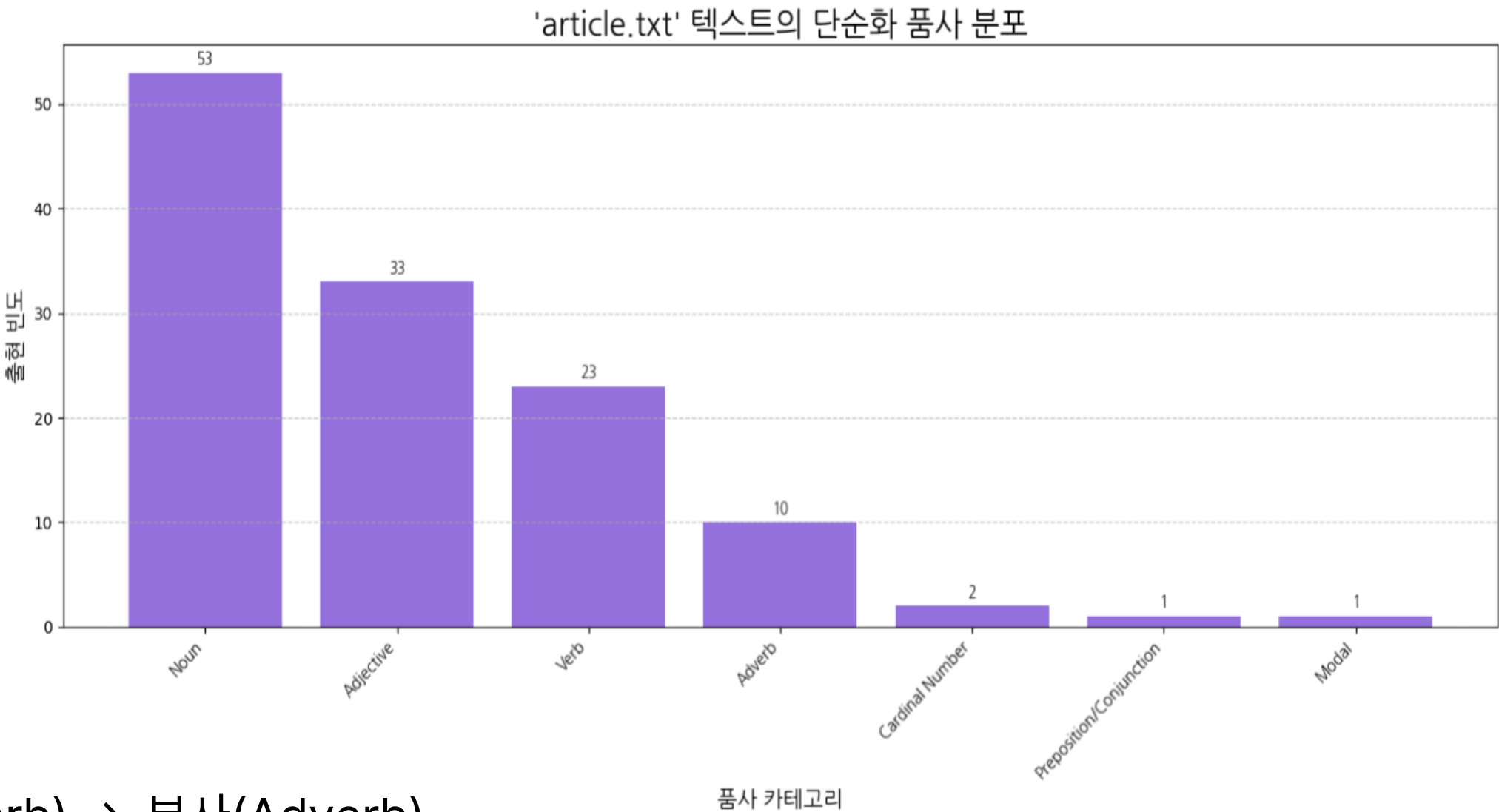
=====

'article.txt'의 품사 분포 시각화

=====

'article.txt'의 통합된 단순화 품사 카테고리 분포:

Noun: 53
Adjective: 33
Verb: 23
Adverb: 10
Cardinal Number: 2
Preposition/Conjunction: 1
Modal: 1



- 가장 많이 등장하는 품사
- 명사(Noun) → 형용사(Adjective) → 동사(Verb) → 부사(Adverb) ...
- 일반적인 글(기사, 설명문 등)에서 명사가 가장 많이 출현하는 것은 흔한 현상임 세상의 대상, 개념 등을 지칭하기 때문



언어학적 분석



idiom 단어 길이 분포 결과

=====

모든 Idiom 텍스트의 단어 길이 분포 시각화

=====

모든 Idiom 텍스트의 평균 단어 길이: 4.96

단어 길이별 빈도수:

길이 3: 3개

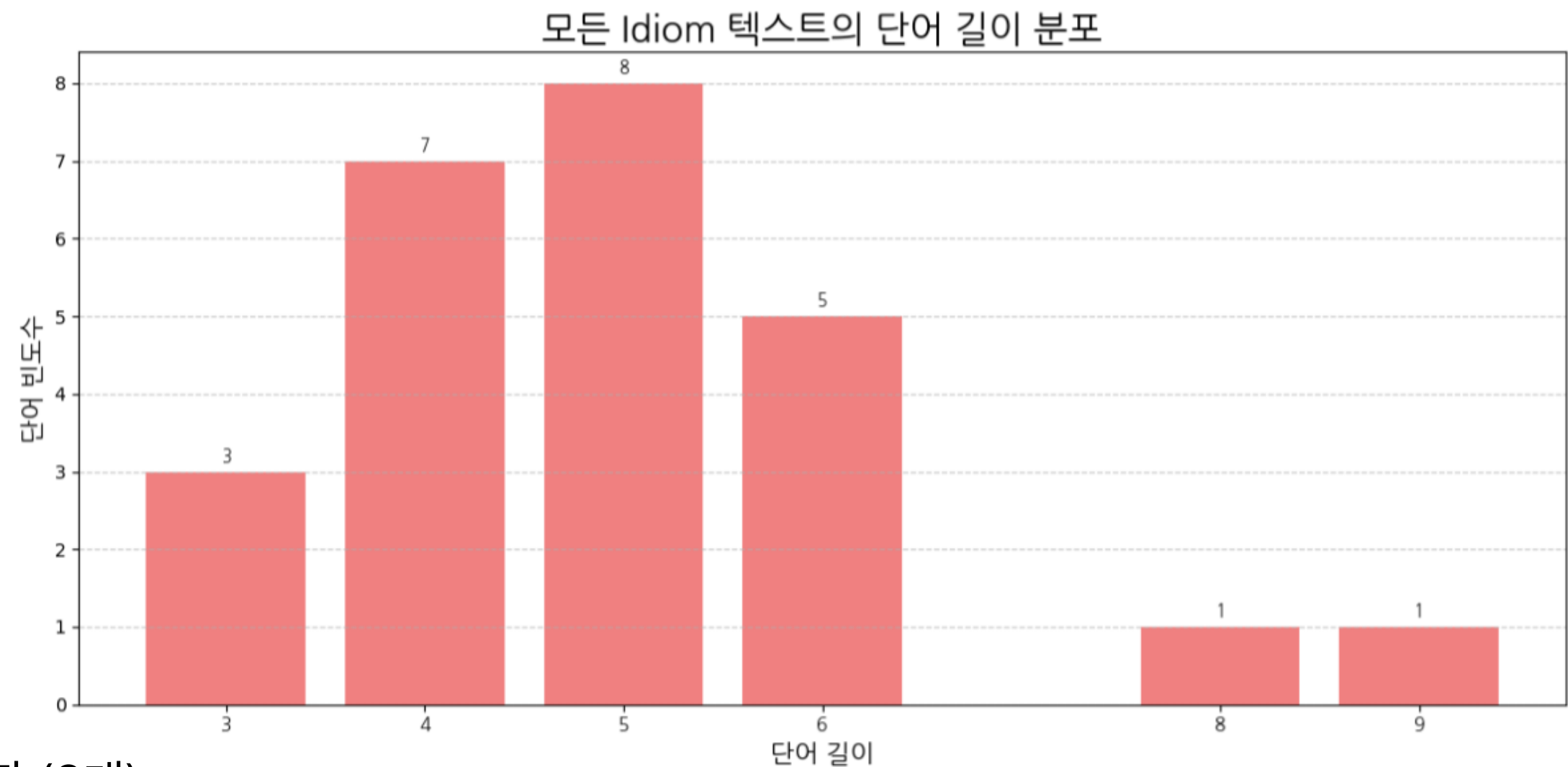
길이 4: 7개

길이 5: 8개

길이 6: 5개

길이 8: 1개

길이 9: 1개



- 가장 빈번하게 나타나는 단어 길이 - 5글자 (8개)
- 전반적으로 짧은 단어(3~6글자) 위주로 구성되어 있음
- 분포가 4~6글자 사이에 집중되어 있으며, 특히 5글자에서 정점
- 매우 긴 단어는 거의 등장 X



언어학적 분석



article 단어 길이 분포 결과

=====

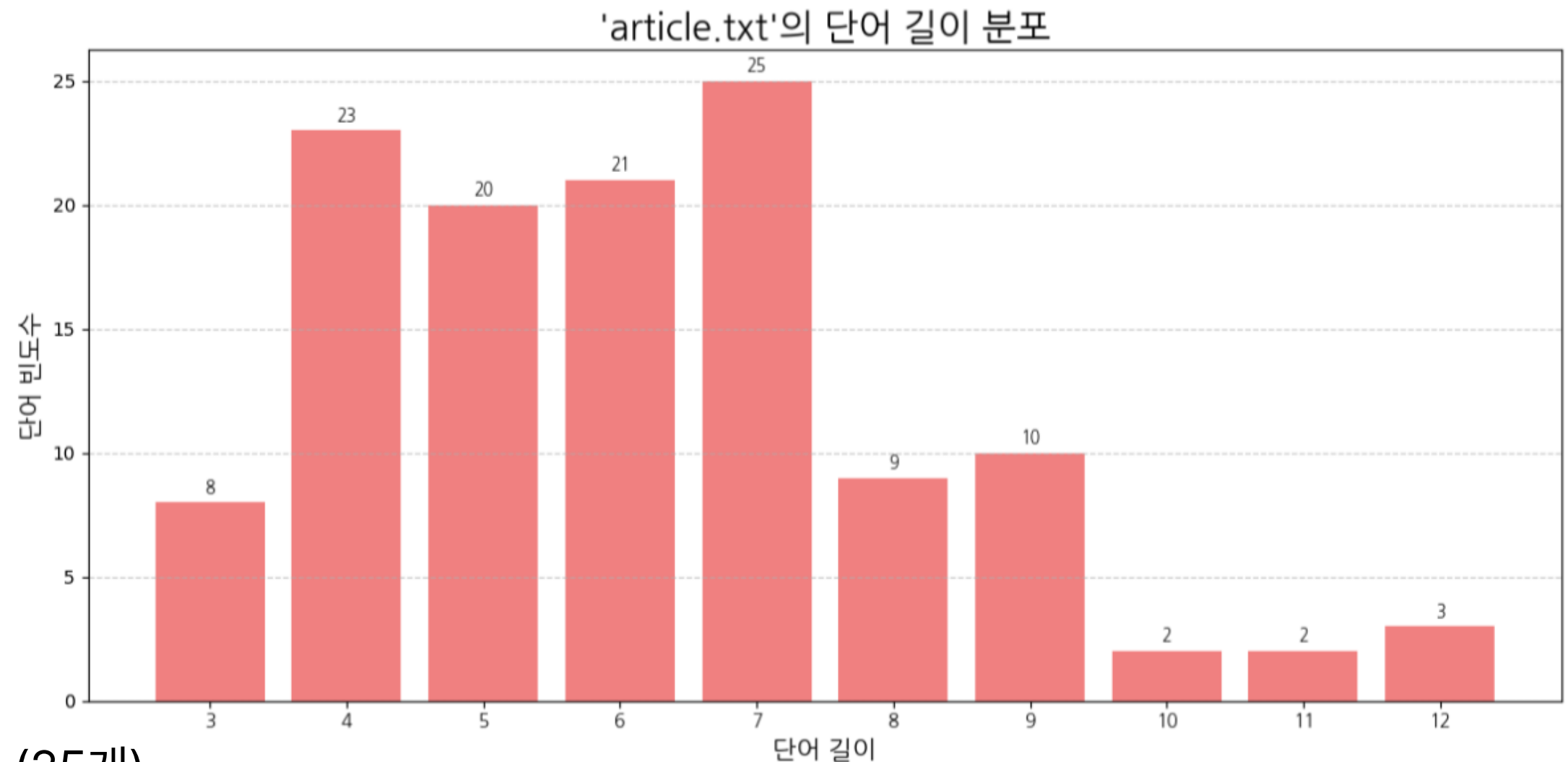
'article.txt'의 단어 길이 분포 시각화

=====

'article.txt'의 평균 단어 길이: 6.15

단어 길이별 빈도수:

길이 3: 8개
길이 4: 23개
길이 5: 20개
길이 6: 21개
길이 7: 25개
길이 8: 9개
길이 9: 10개
길이 10: 2개
길이 11: 2개
길이 12: 3개



- 가장 빈번하게 나타나는 단어 길이 - 7글자 (25개)
- Idiom 텍스트에 비해 단어 길이의 범위가 넓고, 비교적 긴 단어들도 자주 사용
- Idiom 텍스트보다 다양한 길이의 단어들이 사용

- 'article.txt'의 평균 단어 길이가 더 깊
→ 'article.txt'가 일반적인 글(기사, 논문, 설명문 등)일 가능성을 시사
→ 일반적인 글은 관용구(idiom)보다 설명적이거나 전문적인 용어를 포함할 가능성이 높아 단어 길이가 길어지는 경향



언어학적 분석



연구 의의

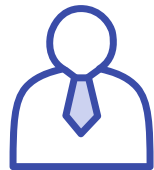
- 텍스트의 성격 파악:
 - 품사 분포는 해당 텍스트의 전반적인 성격(예: 설명문, 문학작품, 구어체 등)을 간접적으로 보여줌
 - 단어 길이 분포는 해당 텍스트의 장르나 성격을 파악하는 데 기초 자료로 활용 가능
- 언어학적 분석 기초 자료:
 - 특정 장르(예: 관용구, 시, 법률 문서 등)의 언어적 특징을 연구할 때 기초 자료로 활용 가능
 - ex) 관용구 분석 결과에서 명사, 형용사, 동사가 주를 이룬다는 점은 관용구의 구조적 특성을 반영



번역 정확도 비교: DeepL vs ChatGPT



- 비교 실험 방법: 1) 파일 업로드 방식 2) 웹사이트 실험



STEP 1 환경 설정 및 데이터 준비

실험에 필요한 라이브러리 설치
API 키(DeepL, ChatGPT) 입력
분석할 데이터 불러오기



STEP 2 핵심 기능 함수 정의

DeepL, ChatGPT를 사용해
양방향 번역 기능 정의
유사도 측정하는 함수 정의



STEP 3 CSV 파일 역번역 실험 실행

DeepL / ChatGPT 번역, 역번역
유사도 계산
결과 출력 및 저장



STEP 4 실시간 비교를 위한 웹 구현

번역 → 역번역 → 유사도 측정
결과를 테이블과 그래프로 시각화
웹 기반 인터페이스 제공



번역 정확도 비교: DeepL vs ChatGPT



STEP 1. 환경 설정 및 데이터 준비

- 실험에 필요한 라이브러리 설치
 - !pip install : 라이브러리 설치
- API 키 (DeepL, ChatGPT) 입력
 - DEEPL_API_KEY / OPENAI_API_KEY = getpass.getpass : API 키 보안 입력
 - import getpass : API 키 숨김 입력
- 분석할 데이터 불러오기
 - files.upload()로 사용자의 컴퓨터에 있는 CSV 파일을 업로드

* API (Application Programming Interface)
- 서로 다른 소프트웨어나 시스템이 서로 소통할 수 있게 해주는 통로



번역 정확도 비교: DeepL vs ChatGPT



- 이용한 텍스트 자료
 - 관용표현 : 기사에서 발췌
 - 일반문장 : 분석에 사용된 관용표현 포함 기사 발췌

Bavarian Football Works. (2024, September 8). Germany faces tough Nations League draw with Hungary

type	sentence																			
idiom	Backed into a corner, Kenny's Ireland need to strike out with purpose in Baku.																			
idiom	Hickman seniors cross the finish line for the last time.																			
idiom	It didn't take long for Pete Crow-Armstrong (PCA) to be a front runner for the National League MVP award when considering he plays for the Chicago Cubs in the MLB 2025 season as a dark horse.																			
idiom	The US will be left to rue their first changeover, with Christian Coleman and Kenny Bednarek essentially running into one another as they attempted to pass the baton.																			
idiom	Lazio flew out of the traps against Inter in an attempt to gain an early advantage, after they failed to score for the first time in 10 league games when they were held to a goalless draw by Torino last time out.																			
normal	Germany's first goal owed a lot to ?lkay G?ndo?an and Jamal Musiala's individual quality.																			
normal	The way the fluid attack consistently created space with enterprising runs and intelligent movement flummoxed the Hungary defense time and again.																			
normal	Germany looked like a much more finished product.																			
normal	The team finally looked more like a Nagelsmann team, with the back three in possession and slow, patient build up play leading to quick, vertical passes in the final third.																			
normal	Comparing Germany's win over Hungary in the EUROs to today's match still demonstrates the progress made under Julian Nagelsmann.																			



번역 정확도 비교: DeepL vs ChatGPT



STEP 2. 핵심 기능 함수 정의

- DeepL, ChatGPT를 사용해 양방향 번역 기능 정의
 - translate_deepl() : 입력된 텍스트를 DeepL API를 통해 번역
 - translate_chatgpt(): OpenAI의 'gpt-3.5-turbo' 모델을 이용해 번역을 수행
- 유사도 측정하는 함수 정의
 - 코사인 유사도: 문장의 의미 유사성을 벡터 각도로 측정
 - > 문장의 전체적인 의미와 맥락이 얼마나 비슷한지 측정
 - BERTScore: 문장 간 의미 차이를 BERT 모델로 정밀하게 비교
 - > 단어 수준의 의미적 유사도를 정밀하게 측정



번역 정확도 비교: DeepL vs ChatGPT



STEP 3. CSV 파일 역번역 실험 실행

- DeepL / ChatGPT 번역, 역번역
 - DeepL과 ChatGPT를 이용해 영어 → 한국어 → 영어 순서로 역번역을 수행
- 유사도 계산
 - 원본 영어 문장과 역번역된 영어 문장 간의 코사인 유사도와 BERTScore를 각각 계산
- 결과 출력 및 저장
 - 원본, 역번역 결과, 계산된 유사도 점수들을 깔끔하게 출력하고 'results' 리스트에 저장



번역 정확도 비교: DeepL vs ChatGPT



STEP 4. 실시간 비교를 위한 웹 구현

- 번역 → 역번역 → 유사도 측정 (step3과 동일)
 - 추가로 가중 종합 점수를 계산 (코사인 유사도에 30%, BERTScore에 70%의 가중치를 부여하여 두 점수를 종합)
- 웹 기반 인터페이스 제공 (Gradio 이용)
- 결과를 테이블과 그래프로 시각화
 - `matplotlib`을 이용해 두 번역기의 코사인 유사도 & BERTScore를 나란히 보여주는 막대그래프 생성
 - 모든 결과(번역 내용, 점수)를 보기 좋은 HTML 표 형식으로 생성



번역 정확도 비교: DeepL vs ChatGPT



결과 설명

- 점수 상 우위: 전반적으로 DeepL이 코사인 유사도와 BERTScore에서 더 높은 점수를 받는 경향
- DeepL
 - 원문의 구조와 뉘앙스를 최대한 그대로 유지하려는 직역에 가까운 스타일을 보임
 - 매우 정확하고 안정적이지만, 때로는 한국어 사용자에게 약간 어색하거나 딱딱하게 느껴질 수 있음

	deepL			chatgpt		
	코사인	bertscore	가중종합	코사인	bertscore	가중종합
idiom1	0.9321	0.9803	0.9658	0.7926	0.874	0.8493
idiom2	0.9499	0.9222	0.9305	0.5722	0.9429	0.8317
idiom3	0.9085	0.9429	0.9326	0.9401	0.9283	0.9369
idiom4	0.8	0.9538	0.9077	0.6553	0.939	0.8798
idiom5	0.656	0.9165	0.8383	0.743	0.9133	0.8364
normal1	0.6691	0.8864	0.8212	0.7514	0.9178	0.8677
normal2	0.8569	0.9481	0.9208	0.8536	0.9358	0.9096
normal3	1	1	1	0.9863	0.9541	0.9637
normal4	0.9697	0.976	0.9741	0.9015	0.9202	0.9351
normal5	0.7613	0.9551	0.8969	0.9243	0.9649	0.9498
평균	0.85035	0.94813	0.91879	0.81203	0.92903	0.896



번역 정확도 비교: DeepL vs ChatGPT



결과 설명

- ChatGPT
 - 원문의 의미를 파악한 후, 한국어에 더 자연스럽게 맥락에 맞는 표현으로 재구성하는 의역 또는 초월 번역에 가까운 스타일을 보임
 - 이 때문에 번역문이 매우 자연스럽게, 원문의 단어나 구조와는 다소 달라져 코사인 유사도 등의 점수는 낮게 나올 수 있음

	deepL			chatgpt		
	코사인	bertscore	가중종합	코사인	bertscore	가중종합
idiom1	0.9321	0.9803	0.9658	0.7926	0.874	0.8493
idiom2	0.9499	0.9222	0.9305	0.5722	0.9429	0.8317
idiom3	0.9085	0.9429	0.9326	0.9401	0.9283	0.9369
idiom4	0.8	0.9538	0.9077	0.6553	0.939	0.8798
idiom5	0.656	0.9165	0.8383	0.743	0.9133	0.8364
normal1	0.6691	0.8864	0.8212	0.7514	0.9178	0.8677
normal2	0.8569	0.9481	0.9208	0.8536	0.9358	0.9096
normal3	1	1	1	0.9863	0.9541	0.9637
normal4	0.9697	0.976	0.9741	0.9015	0.9202	0.9351
normal5	0.7613	0.9551	0.8969	0.9243	0.9649	0.9498
평균	0.85035	0.94813	0.91879	0.81203	0.92903	0.896



번역 정확도 비교: DeepL vs ChatGPT



번역기별 장단점 및 연구 의의

- 번역기별 장단점

- DeepL

- 장점

- 높은 정확성과 충실성: 원문의 단어와 문장 구조를 거의 그대로 가져와 번역하므로, 정보의 누락이나 왜곡이 적음
- 안정적인 성능: 어떤 종류의 문장이든 일관성 있고 예측 가능한 번역 품질을 보여줌
복잡하고 긴 문장에서도 구조를 잘 유지

- 단점

- 문맥적 유연성 부족: 때로는 문맥상의 숨은 의미를 파악하지 못하고 표면적인 의미로만 번역할 수 있음
- 어색한 표현: 직역 스타일 때문에 한국어로는 잘 쓰이지 않는 딱딱한 문체가 될 수 있음



번역 정확도 비교: DeepL vs ChatGPT



번역기별 장단점 및 연구 의의

- 번역기별 장단점

- ChatGPT

- 장점

- 뛰어난 문맥 파악 및 자연스러움: 문장의 숨은 의미나 관용적인 표현을 잘 파악하여 매우 자연스러운 한국어 문장을 만들어냄

- 창의적인 표현 구사: 원문을 단순히 옮기는 것을 넘어, 더 적절하고 세련된 표현으로 재창조하는 능력이 있음

- 단점

- 낮은 평가 점수: 번역이 자연스러워도 원문과 구조가 달라지기 때문에, 코사인 유사도 같은 자동 평가 지표에서는 오히려 낮은 점수를 받을 수 있음. 이는 '점수'가 '번역 품질'과 항상 일치하지는 않음을 보여줌.



번역 정확도 비교: DeepL vs ChatGPT



번역기별 장단점 및 연구 의의

- 연구 의의
 - AI 번역기의 성능 비교 지표 정립
 - 코사인 유사도, BERTScore와 같은 정량적 평가 지표를 활용해 번역 품질을 비교
 - 사람 주관에 의존하지 않고 객관적인 수치로 번역기 성능 평가 가능
 - 관용표현 번역의 중요성 조명
 - 스포츠 기사 등에서 자주 등장하는 관용적 표현은 직역하면 의미가 왜곡될 수 있음
 - 이 실험을 통해 표면적 의미와 실제 의미 간의 간극 측정 가능
 - 교육 및 번역 시스템 개발에 활용 가능
 - 어떤 번역기가 어떤 문장 유형(관용 vs 일반)에 강한지 파악 → 번역기 선택 기준 제시
 - 향후 관용 표현에 특화된 하이브리드 번역 시스템 개발 근거 제공



번역 정확도 비교: DeepL vs ChatGPT



번역기별 장단점 및 연구 의의

- 연구 의의
 - 실험의 대중화 및 재현성 향상
 - 누구나 웹사이트에 접속해 문장을 직접 입력하고 실험 결과 확인 가능
 - 연구 결과를 단순히 논문으로 제시하는 것을 넘어, 대중·비전문가도 실험을 직접 체험할 수 있게 함
 - 인공지능 번역의 발전 방향
 - 과거의 기계 번역이 단어 대 단어의 치환에 그쳤다면, 이제는 문맥 전체를 이해하고 문화적 배경까지 고려하여 번역하는 수준으로 발전하고 있음을 보여줌
 - 이는 앞으로 번역의 패러다임이 어떻게 변할지를 예측하게 해줌
 - 사용자 맞춤형 번역 실험 환경 제공
 - 일반 기사, 관용 표현 등 연구자가 미리 준비한 데이터뿐만 아니라, 사용자가 직접 입력한 문장에 대해서도 번역 성능 비교 가능 → 사용자 중심의 인터랙티브 실험 환경



출처



- 번역 정확도 비교에서 이용한 관용표현 자료

- Backed into a corner, Kenny's Ireland need to strike out with purpose in Baku.

- <https://www.irishtimes.com/sport/soccer/international/backed-into-a-corner-kenny-s-ireland-need-to-strike-out-with-purpose-in-baku-1.4695457>

- Hickman seniors cross the finish line for the last time.

- <https://purpleandgoldnews.com/3687/sports/hickman-seniors-cross-the-finish-line-for-the-last-time/>

- It didn't take long for Pete Crow-Armstrong (PCA) to be a front runner for the National League MVP award when considering he plays for the Chicago Cubs in the MLB 2025 season as a dark horse.

- <https://timesofindia.indiatimes.com/sports/mlb/news/why-pete-crow-armstrong-could-be-the-dark-horse-mvp-of-2025-mlb-season/articleshow/121524193.cms>

- The US will be left to rue their first changeover, with Christian Coleman and Kenny Bednarek essentially running into one another as they attempted to pass the baton.

- <https://www.bbc.com/sport/olympics/articles/c623lv6rg56o>

- Lazio flew out of the traps against Inter in an attempt to gain an early advantage, after they failed to score for the first time in 10 league games when they were held to a goalless draw by Torino last time out.

- https://www.espn.com/soccer/report/_/gameId/644645

