

초등 4학년 수학 강의 검색 챗봇 개발

빅데이터 9기 | 정승연, 박만혜, 조아라, 신유라

목 차

01

개요

02

개발 환경

03

개발 프로세스

04

데이터 준비

05

챗봇 서비스 구현

06

성능 평가

07

개선점

08

교육 서비스 응용

개요

개발 목표

수학 강의 정보 제공 챗봇 개발(초등학교 4학년 대상)

- 학습 키워드를 포함한 질문 입력 시, 관련 강의 썸네일 및 요약 정보 출력

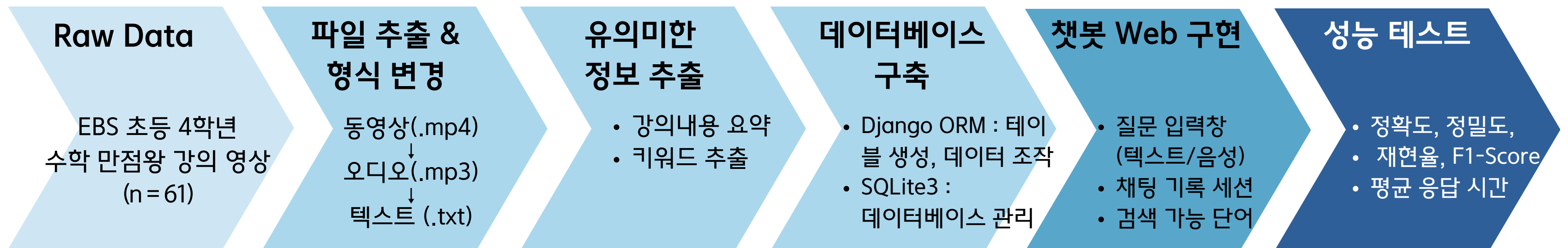
프로젝트 기간

2024년 09월 12일 - 2024년 09월 20일

개발 환경

언어 및 라이브러리	버전	라이선스
Python	3.9.19	Python Software Foundation License
Pandas	2.2.2	BSD 3-Clause License
Django	4.2.16	BSD 3-Clause License
Transformers	4.44.2	Apache License 2.0
OpenAI-Whisper	0.28.0	MIT License
KoNLPy	0.6.0	GPL v3 License
Selenium	4.23.1	Apache License 2.0

개발 프로세스



개발 프로세스

팀 프로젝트 | 2024. 09. 20

상세 프로세스

음성-텍스트 변환

- Whisper(ver. medium) 모델로 수학 강의 영상을 음성 텍스트로 변환
- 강의 내용에 대한 요약 및 키워드 추출 후, 데이터베이스에 저장

문장에서 키워드 추출

- 한국어 형태소 분석기 KoNLPy 사용
- 입력된 질문에서 명사와 형용사 키워드를 추출, 이를 데이터베이스와 매칭

입력된 문장 개선

- LLM(Large Language Model) 사용
- 학생들의 질문 자동 개선 및 오타 수정 후, 정확한 자료 추천

정보 검색형 챗봇 구현

- Django 프레임워크를 사용하여 학생들의 질문 분석
- 관련된 강의 정보를 제공하는 정보 검색형 챗봇을 개발

데이터 준비

데이터 베이스 생성

- Django ORM을 통해 DB(Data Base) Table 생성

데이터 베이스 관리

- SQLite3로 경량화된 DB 테이블 관리
- 소규모 프로젝트에서 쉽게 데이터를 저장 및 처리

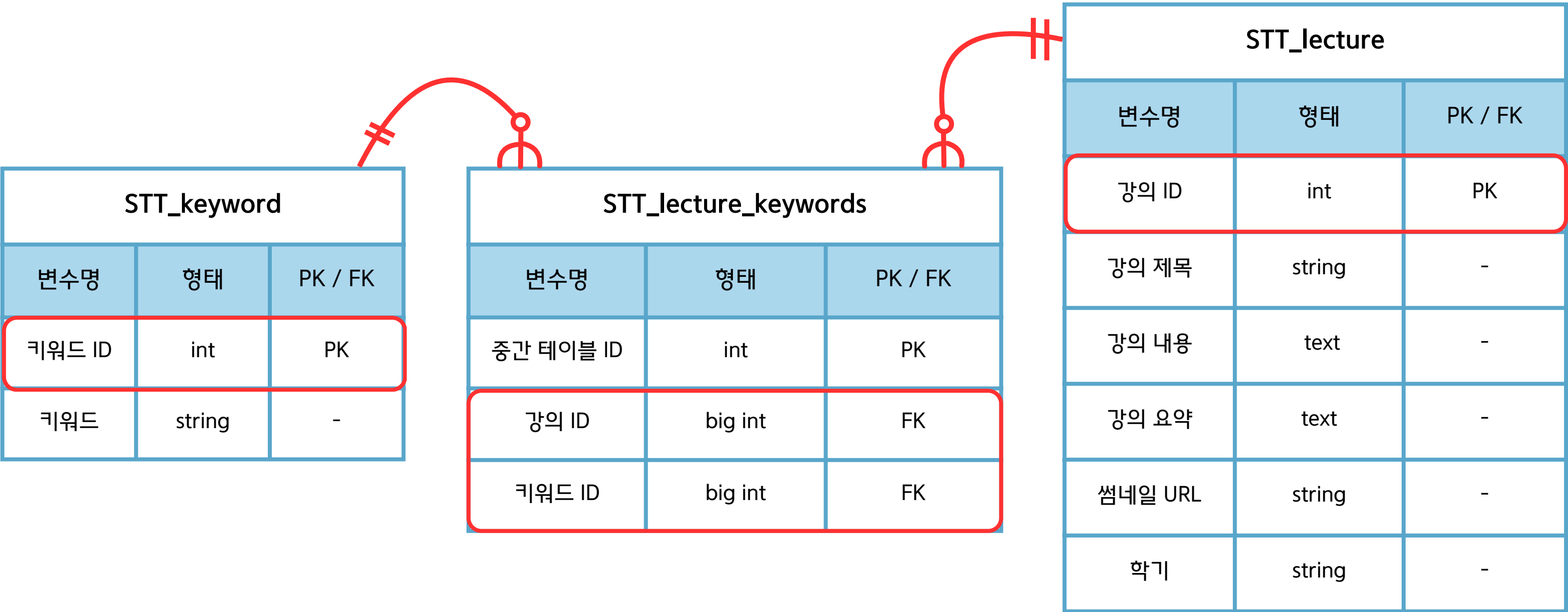
데이터 테이블

- STT_keyword: 키워드 정보
- STT_lecture: 강의 제목, 강의 내용, 강의 요약, 썸네일 URL, 학기 정보
- STT_lecture_keywords: 두 테이블을 연결

데이터 준비

팀 프로젝트 | 2024. 09. 20

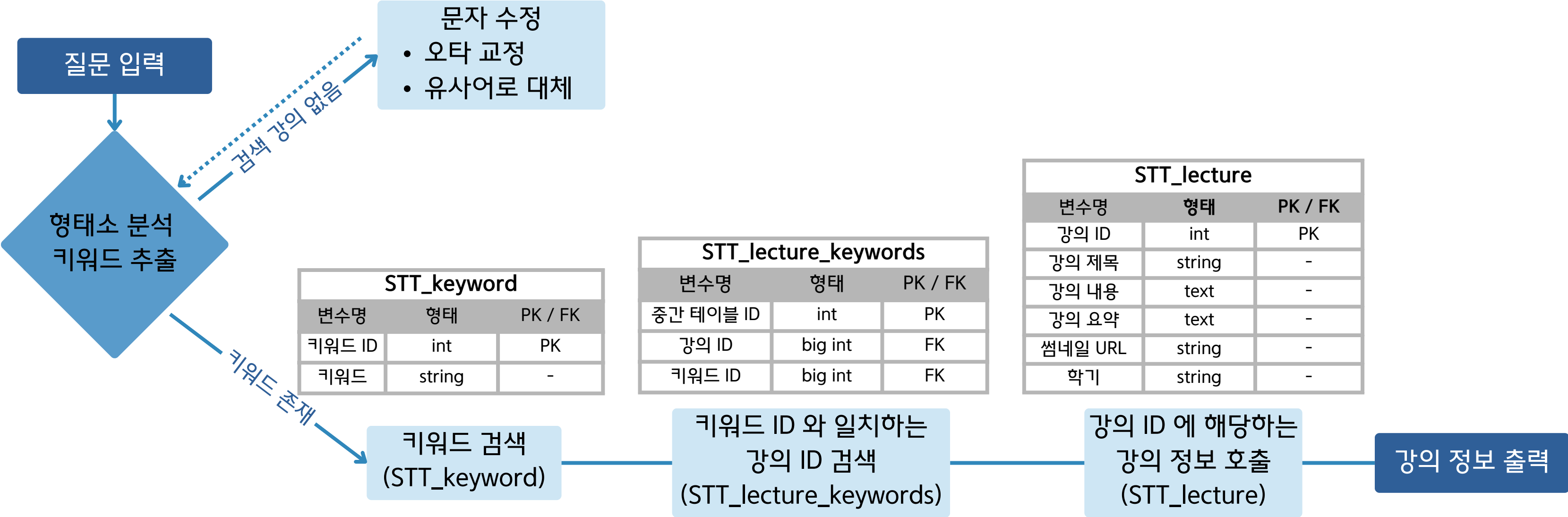
생성된 데이터 테이블



챗봇 서비스 구현

팀 프로젝트 | 2024. 09. 20

챗봇 작동 방식



챗봇 서비스 구현

팀 프로젝트 | 2024. 09. 20

챗봇 화면

질문 입력창

○ →

질문하기

음성 입력

○ ← 텍스트/음성 입력 가능

답변 출력

○ →

챗봇의 답변

채팅 기록

○ ← 채팅 기록 자동 저장

검색 가능 단어 표시

○ →

검색 가능 단어

큰 수, 곱셈, 사각형, 꺾은선그래프, 소수, 분수, 다각형, 평면도형, 막대그래프, 각도, 규칙 찾기, 삼각형, 나눗셈

채팅 기록 삭제

○ ← 채팅 기록 삭제 가능

성능 평가

팀 프로젝트 | 2024. 09. 20

평가 방법

테스트 데이터셋 생성
(n = 100)

- ① SQL 쿼리문을 통해 키워드, 강의명 데이터 호출
- ② 키워드 - 강의명 데이터 매핑
- ③ 키워드별 질문 랜덤 생성 & 관련된 강의 리스트를 기대 답변으로 설정
- ④ 질문과 답변으로 구성된 테스트 데이터셋 구축

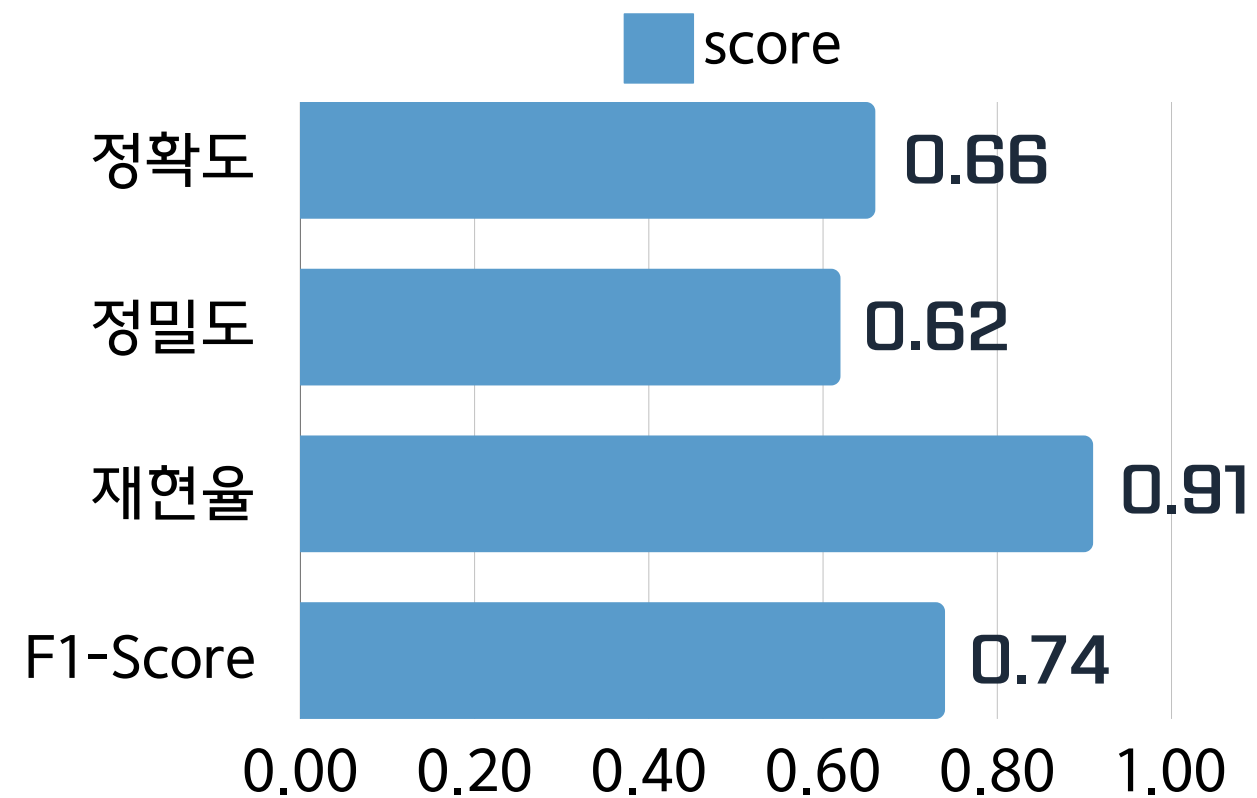
평가 항목

- 정확도, 정밀도, 재현율, F1-Score
- 평균 응답 시간

성능 평가

팀 프로젝트 | 2024. 09. 20

평가 결과



평균 응답 시간 : 0.13 초

낮은 정밀도

키워드 추출 시 복합 명사를 단일 명사로 분리하여 복합 명사와 단일 명사 키워드 강의 모두 추출 → 거짓 양성 비율 증가

높은 재현율

실제 찾고자 하는 강의를 포함하여 여러 강의 검색

빠른 응답시간

고객이 높은 만족도를 보이는 평균 반응 속도인 2~5초보다 빠름

개선점

음성 인식 기능 개선

다른 브라우저와의
호환성

음성 인식 실패 시
텍스트 입력 전환 유도

검색 기능 강화

유사한 키워드 포함
강의 검색 지원

자연어 처리 모델로
더욱 정확한 답변 제공

국제화 및 다국어 지원

다국어 번역 지원으로
외국인 도움

보안강화

사용자 인증 기능으로
개인화 서비스 제공

교육 서비스 응용

개인 맞춤형 학습 경로 제공

- 학생의 질문 패턴과 학습 성향 분석
- 학생의 학습 수준과 목표에 맞는 강의 추천

교육용 챗봇

- 학습 중 질문 입력 시 실시간으로 관련 강의 및 자료 탐색 후 정보 제공

감사합니다