

초등 4학년 수학 챗봇 개발: 강의 검색 성능 분석

[빅데이터 9기] 정승연, 신유라, 박만혜, 조아라

[목차]

1. 요약
2. 서론
3. 개발 환경
4. 개발 프로세스
5. 데이터 준비
6. 챗봇 서비스 구현
7. 결과
8. 분석 및 토의
9. 결론
10. 소스코드
11. 참고문헌

1. 요약

본 보고서는 초등학교 4학년 수학 교과과정을 안내하는 정보 검색형 챗봇의 개발과정과 평가 내용을 다룬다. EBS에서 제공하는 71개의 초등학교 4학년 만점왕 수학 영상데이터를 활용했으며 트랜스포머 아키텍처를 활용한 음성-텍스트 변환 시스템을 비롯하여 데이터베이스, Django 프레임워크를 활용하여 정보 검색형 챗봇 모델을 구현하였다. 성능 평가를 실시한 결과, 정확도 0.87, 정밀도 1.00, 재현율 0.87, F1-Score 0.93, 평균 응답 시간 0.17초의 성능을 나타내었다. 본 모델을 통해 영상 및 음성의 비정형자료로부터 얻은 정보는 고객 서비스 제공에 활용될 가능성이 있음을 제시한다.

2. 서론

1) 배경

학습용 챗봇은 학습자의 요구에 맞춘 개인 맞춤형 학습을 지원하는 도구로, 다양한 플랫폼에서 활용되고 있다. 현재 상용화된 유사 기술로는 Coursera와 Moodle에서 제공하는 챗봇이 있으며, 이 챗봇들은 사용자가 입력한 키워드와 학습 활동 기록을 활용하여 강의를 추천하는 시스템을 제공한다. 이러한 챗봇은 학습자의 필요에 맞는 강의를 신속하게 찾아주며, 특히 자기주도 학습에 있어 중요한 역할을 한다.

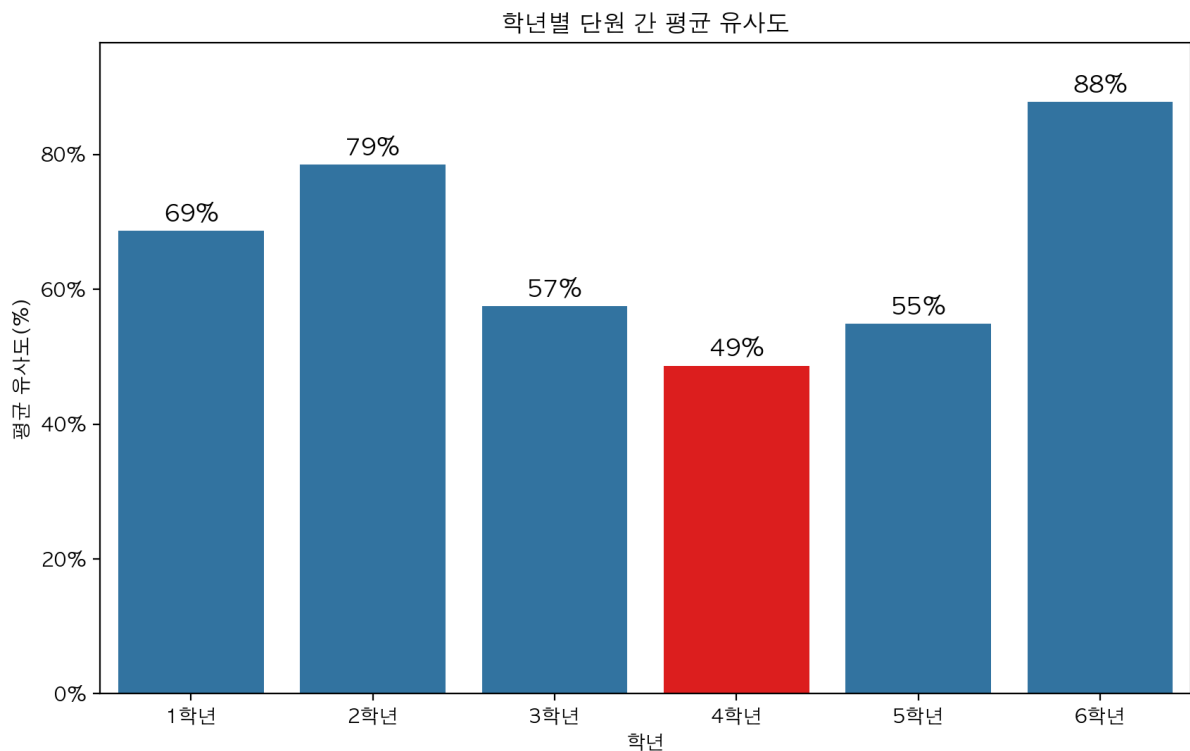
자기주도형 학습이 중요한 시기에 챗봇은 학생들이 학습 계획을 세우고 스스로 학습을 진행하는 데 실질적인 도움을 제공한다. 학생들은 필요한 정보를 챗봇에 입력하여 키워드 추출을 통해 관련된 학습 자료를 즉시 제공받을 수 있으며, 이를 통해 적합한 학습 자료를 손쉽게 탐색하고 학습의 효율성을 높일 수 있다. 챗봇은 이러한 방식으로 학생들이 스스로 학습 목표를 설정하고 실행할 수 있도록 돕는다.

2) 기획의도

본 프로젝트는 초등학교 4학년 수학 강의 영상 데이터를 활용해 강의 정보 검색형 챗봇을 구현하는 것을 목표로 한다. 4학년은 자기주도 학습이 본격적으로 시작되는 시기이며, 피아제의 인지 발달 이론¹에 따르면 논리적 사고와 구조적 과목에 대한 이해가 발달하는 단계이기도 하다.

¹ 김혜진 (2009). 「피아제(J. Piaget) 인지발달이론의 교육적 시사점에 비추어 본 초등학교 수학교과서의 적절성 분석 : 5, 6, 7차를 중심으로」. 석사학위논문, 창원대학교 대학원, p.6~8.

추가적으로, 각 학년별 수학 강의 데이터²의 유사도 분석을 진행하였다. 총 352개의 EBS 초등수학 만점왕(2024) 강의 데이터를 스크래핑하여 동일 학년 내에서 강의별로 단원 제목, 강의 내용과 요약들을 묶은 후 각 그룹이 서로 얼마나 유사한지 TF-IDF로 벡터화한 뒤 코사인 유사도를 계산하였다. 그 결과, 4학년의 유사도가 49%로 가장 낮게 나타났고, 이는 곧 4학년 수학과목의 단원별로 주제가 가장 명확하게 나뉘어 있다는 것을 의미한다. 이 분석을 통해, 챗봇에 특정 키워드를 입력했을 때 관련 강의가 과도하게 많이 출력되지 않고, 사용자가 원하는 강의를 보다 정확하게 찾아서 제공할 수 있다는 가능성을 확인하였다.



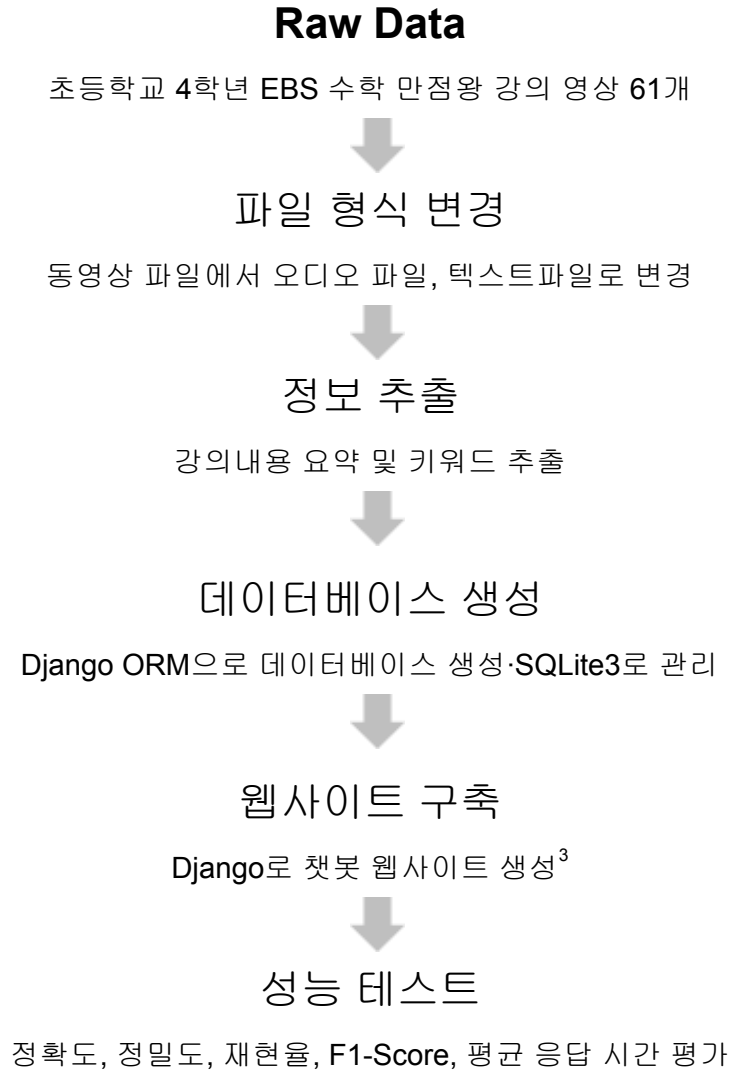
이에 최종적으로 초등학교 4학년 수학 학습자를 대상으로 설정하여 챗봇 서비스를 구현하였다.

² 유사도 분석에 사용된 수학 강의 데이터에는 강의 제목, 수업 내용, 키워드가 포함되어 있다

3. 개발 환경

언어 및 라이브러리	버전	라이선스
Python	3.9.19	Python Software Foundation License
Pandas	2.2.2	BSD 3-Clause License
FFmpeg	7.0.2	GNU Lesser General Public License (LGPL)
MoviePy	1.0.3	MIT License
Django	4.2.16	BSD 3-Clause License
Selenium	4.23.1	Apache License 2.0
Transformers	4.44.2	Apache License 2.0
OpenAI	0.28.0	MIT License
PyTorch	2.4.1+cpu	BSD 3-Clause License
SpaCy	3.7.6	MIT License
PyMySQL	1.1.1	MIT License
KoNLPy	0.6.0	GPL v3 License
Pytz	2024.2	MIT License
Googletrans	4.0.0rc1	MIT License
Requests	2.32.2	Apache License 2.0

4. 개발 프로세스



³ 크롬에 최적화된 챗봇을 구현하고자 함

5. 데이터 준비

- 사용 데이터

- 초등학교 4학년 EBS 수학 만점왕 수업 영상 데이터 61개⁴

- 데이터 형식 변환

1) 동영상 파일을 오디오 파일로 변환

- 텍스트 변환 작업 시, 작업 속도를 줄이기 위해 동영상에 비해 용량이 작은 오디오 파일로 변환
- MoviePy 라이브러리의 VideoFileClip 클래스 이용 : mp4 동영상 파일에서 오디오 트랙을 분리 → mp3 파일로 저장

2) 오디오 파일을 텍스트 파일로 변환

- 문장의 완성도와 속도 및 비용을 고려하여 Transformers 라이브러리 Whisper 모델 medium 버전 사용
- mp3 오디오 파일을 텍스트 파일로 저장

3) 텍스트 파일 요약 및 키워드 추출

- OpenAI의 API를 사용 : 텍스트 파일로 저장된 전체 수업내용 요약 및 키워드 추출
- API 요청사항
 - 불필요하고 중복된 내용 제거
 - 수업 내용은 최대 글자 수 600자로 설정
 - 키워드에 숫자가 포함되면 삭제
 - 요약은 5개의 항목으로
 - 첫인사와 마지막 인사 제거

4) EBS 웹 페이지의 강의설명과 키워드 추출

- Chrome 웹 드라이버 사용 : EBS 웹 페이지의 각 강의별 강의 제목, 키워드, 수업 내용, 썸네일 URL 스크래핑
- XPATH 사용 : 강의 제목이 포함된 strong 태그, 강의 설명·키워드가 포함된 dt-dd 태그, 썸네일 이미지 URL 추출 → 리스트에 저장

⁴ 총 71개 중 개념설명이 들어있지 않는 강의와 OT 영상 파일을 제외한 61개의 강의 영상 선정

<https://laurencemoroney.com/datasets.html#horses-or-humans-dataset>.

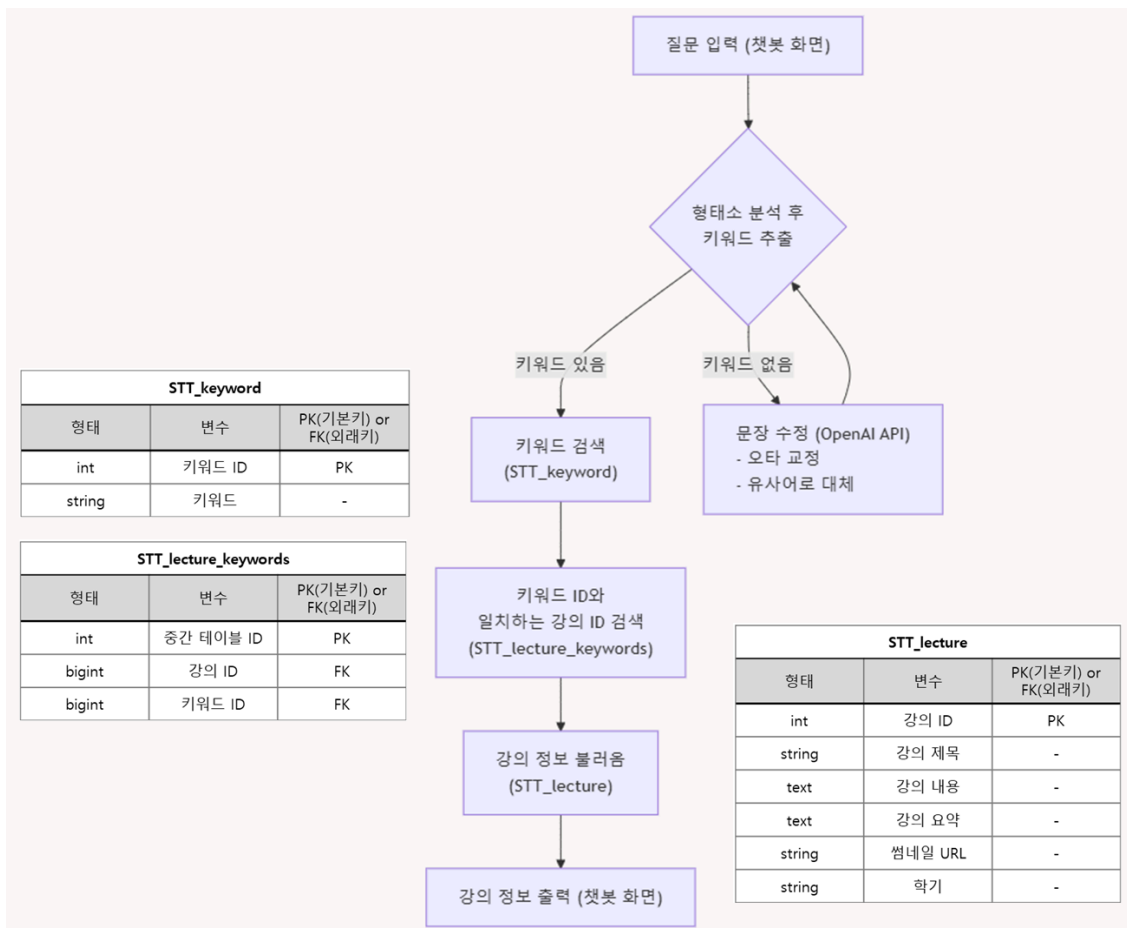
- 저장된 리스트를 데이터프레임으로 변환 → CSV로 최종 저장

5) 3)과4) 내용 통합

- 영상에서 추출된 정보와 웹사이트에 게재된 정보를 취합하여 CSV파일('final_data.csv')로 저장 : 정보 검색 시 답변의 정확도를 높임

6. 챗봇 서비스 구현

- 작동 방식



- 데이터베이스 구축

- Django ORM을 통해 SQLite3와 연결된 데이터베이스를 구현
- 앞서 취합된 CSV파일('final_data.csv')을 불러와 강의와 키워드에 대한 테이블생성 : 강의 제목, 강의 내용, 강의 요약, 썸네일 이미지 URL, 학기, 키워드 정보를 포함
- 두 테이블(STT_keyword, STT_lecture)은 STT_lecture_keywords 테이블을 활용하여

다대다 관계를 갖도록 설정

- 강의 ID와 키워드 ID는 각 테이블의 외래키를 나타냄

- 챗봇 서비스 웹페이지 생성

1) Django 모델 정의

- **Lecture** 클래스

강의 제목, 강의 내용, 강의 요약, 썸네일 이미지 URL, 학기 정보

- **Keyword** 클래스

강의와 관련된 키워드를 저장하며 **Lecture** 클래스와 다대다 관계 구축

2) 질문 처리와 응답 시스템

① 사용자가 질문 입력(POST)

② **KoNLPy** 라이브러리의 **Okt** 형태소 분석기를 사용하여 질문에서 불필요한 단어(불용어, 조사, 구두점 등)를 제거하고 명사·형용사 키워드 추출

③ 데이터베이스에서 키워드와 매칭되는 강의 썸네일 및 요약 정보 추출

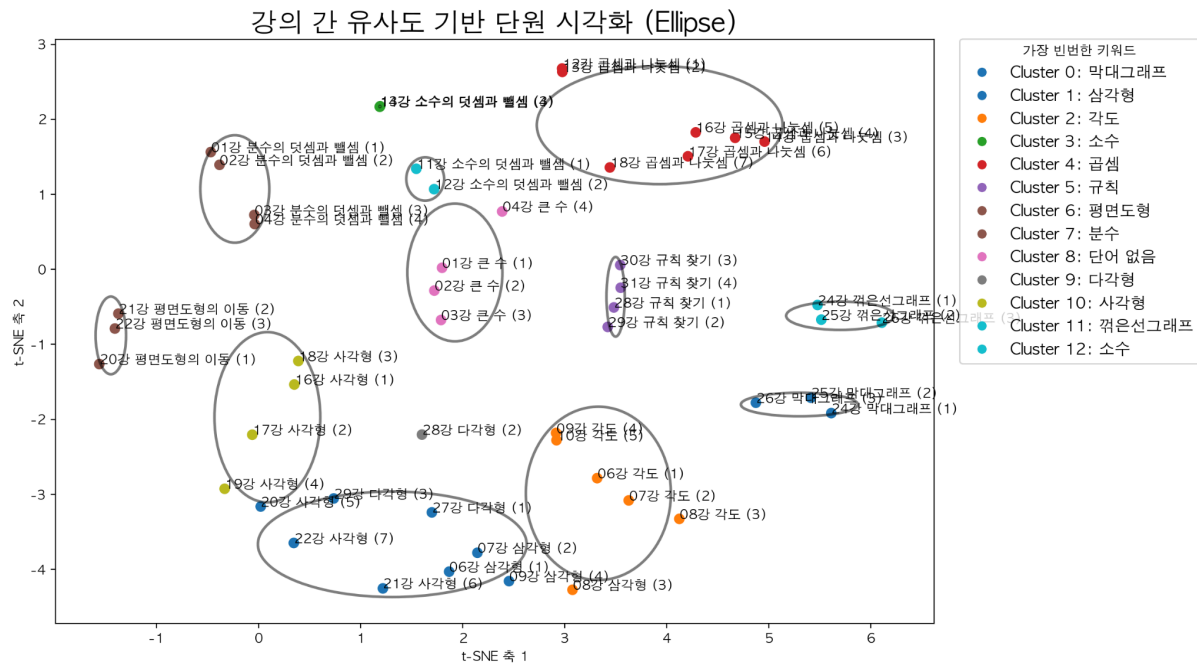
* 강의가 매칭되지 않는 경우, **OpenAI**의 **API**를 사용하여 질문에 포함된 오타 수정 or 유사어 대체 → 다시 ②로 이동(1회)

④ 강의 정보를 템플릿에 전달 → 챗봇 화면에 렌더링

⑤ 대화 기록을 세션에 저장 : 필요 시 삭제하는 기능 포함

3) 템플릿(사용자 인터페이스) 구성

- 검색 가능 단어 챗봇 화면 하단에 제시 : 빈도수 별 클러스터링을 통해 13개의 키워드 선정



- 질문 입력 폼을 통해, 사용자가 텍스트·음성 형태로 질문을 입력(POST)
- 챗봇의 응답은 템플릿에서 동적으로 렌더링
- 서버에서 처리된 데이터는 태그를 사용하여 표시 → JavaScript를 통해 로딩 스피너 및 음성 인식 기능 제공
- 서버에서 반환된 응답은 HTML을 통해 사용자에게 표시

4) 음성 인식 기능

- Chrome에서 지원하는 웹 음성 인식 API인 **webkit Speech Recognition**를 이용하여 음성 입력을 처리
- 입력된 음성을 텍스트로 변환하여 질문 입력란에 자동으로 삽입 후 제출하여 질문에 대한 답을 제공

- 모델 성능 테스트

1) 테스트 데이터셋 구축 (n=100)

- 5가지 질문 유형 설정
 - 단순질문
 - 오타가 포함된 질문
 - 복합 질문
 - 학습자(초등학교 4학년) 수준의 질문

- 구체적 설명을 요구하는 질문
- 각 질문 유형별로 3가지 세부 질문 형식을 설정한 후 키워드-강의명을 랜덤으로 배정
- 복합질문의 일부에는 오타를 추가하여 각 키워드에 대해 강의명 검색이 가능한지 확인

2) 성능 테스트

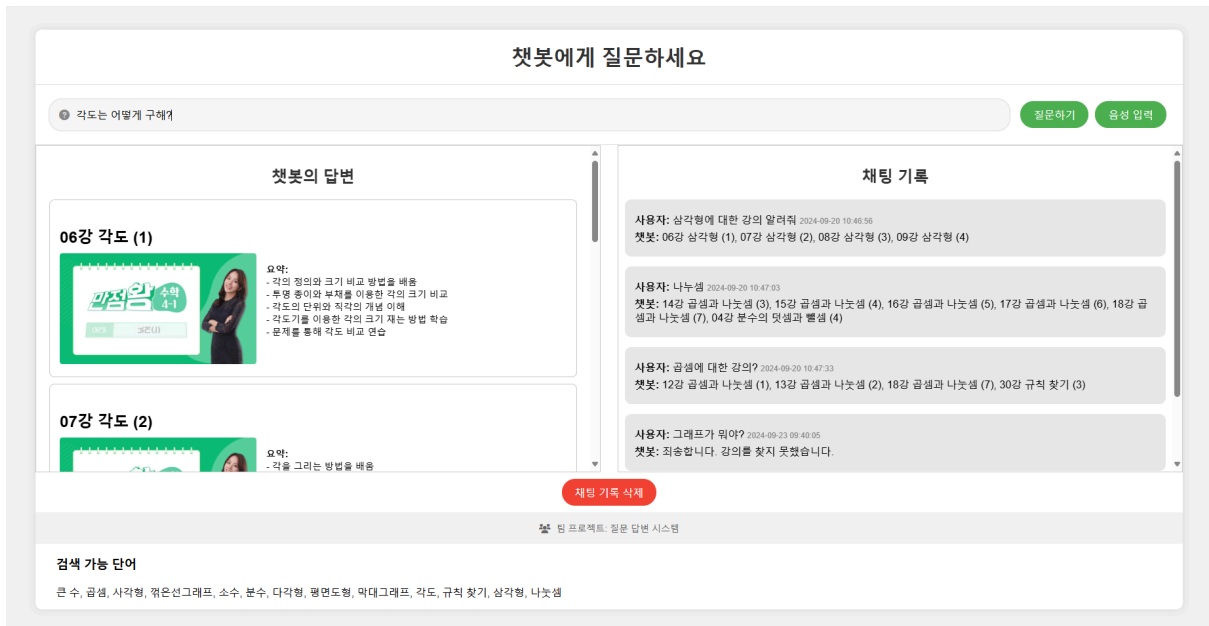
- 정확도, 정밀도, 재현율, F1 스코어, 평균 응답 시간
- 각 질문에 대해 예측된 강의명 중 예상답안이 존재하는 경우 양성으로 판단
- 복합질문은 각 키워드별로 예측된 강의명 중 예상답안이 모두 존재해야 양성으로 판단

7. 결과

1) 챗봇 서비스 구현

▪ 사용자 인터페이스

- 질문 입력창 : 텍스트 및 음성 입력 가능
- 답변 출력창 : 강의가 여러 개일 경우 스크롤 생성
- 채팅 기록 : 세션에 자동 저장되며 필요 시 삭제 가능
- 검색 가능 단어 표시 : 더 정확한 강의 검색을 유도하기 위해 화면에 상시 표시



2) 모델 성능 테스트⁵

▪ 성능 지표 분석

	Score
Accuracy	0.79
Precision	1.00
Recall	0.79
F1-Score	0.88

▪ 평균 응답 시간 : 0.17 초

⁵ 성능 지표 값 및 평균 응답 시간은 소수점 셋째 자리에서 반올림하였다

8. 분석 및 토의

▪ 챗봇 모델의 성능

챗봇 모델은 정확도 0.79, 정밀도 1.00, 재현율 0.79의 성능과 0.17초의 빠른 평균 응답 시간을 보였다. 질문 입력 시 모델이 관련 항목들을 충분히 포괄하여 출력하기 때문에 정밀도가 높게 나왔다. 이는 챗봇이 예측한 강의명에 예상답안이 모두 포함되었음을 나타낸다. 기본적으로 장고 ORM을 사용하지만 강의를 찾지 못한 경우에만 API(OPEN AI)를 호출하여 강의를 검색하도록 처리하였기 때문에 빠른 속도를 유지할 수 있었다. 재현율과 정확도가 다소 떨어지는 것은 검색할 수 있는 모든 강의 중 일부를 놓치고 있기 때문이다. 즉, 예상답안 중 하나가 모두 예측하는 강의에 포함되었으나, 모든 예상답안을 검색하지 못하였다는 것이다. F1 스코어는 0.88로, 정밀도는 높지만 재현율이 낮기 때문에 비교적 낮게 나왔다.

질문에 복합 명사가 포함되는 경우, 키워드 중간에 오타가 포함되거나 자모음이 추가되는 경우, 복합질문에 대해 하나의 키워드에 대한 강의를 검색하지 못하는 경우에 성능이 저하되는 것으로 판단하였다. 일부 복합명사에 대해 단일명사로 분리하여 단일명사에 대한 강의만을 검색하는 경우가 있었다. 또한 키워드 중간에 오타가 있을 때 오타를 교정하지 못하고 한 글자씩 분리하거나 중간에 자모음이 추가되면 그 자모음을 기준으로 키워드를 분리하여 처리하는 경우를 보였다. 복합질문에 오타가 있는 경우에도 비슷한 문제점이 드러났다. 이러한 문제점으로 인해 재현율과 정확도가 다소 떨어진다고 해석할 수 있다.

정밀도가 1.00으로 나온 것은 데이터의 양이 적고 테스트 데이터셋의 질문에 키워드를 직접적으로 추가하여 질문하도록 유도하였기 때문이다. 또한 예측된 여러 강의 중 예상하는 강의 전부가 아닌 하나라도 포함될 경우에 양성으로 판단하는 성능 평가 방식이 영향을 주었다.

▪ 개선점

1) 브라우저 호환성

- 현재 챗봇 서비스는 크롬에 최적화된 서비스 구현을 목적으로 하였으나 파이어폭스, 마이크로소프트 엣지, 오페라와 같은 다른브라우저에서도 작동할 수 있도록 webkit Speech Recognition 외 다양한 음성인식 API나 플러그인을 활용해볼 수 있다.

2) 음성 인식 기능 개선

- 음성 인식이 실패했을 때 사용자에게 알림을 주고 텍스트 입력으로 전환하도록 유도할 수 있다.

3) 검색 기능 강화

- **KoBERT 적용** : 현재 챗봇은 오타 수정과 유사어 대체 기능을 제공하고 있으나, 오타 처리에 큰 문제를 보이고 있다. 이에 **KoBERT** 모델을 적용하여 사용자의 질문 의도를 더 깊이 이해할 수 있도록 개선할 수 있다. **KoBERT**는 한국어에 특화된 **BERT** 모델로, 문장의 맥락을 정확하게 파악하여 더 정확하고 세밀한 답변을 제공할 수 있으며, 복잡한 질문에 대해서도 보다 적합한 강의를 추천하는 검색 기능을 강화할 수 있다

4) 사용자 인증 기능 추가

- 현재 별도의 사용자 인증 기능이 없지만, 로그인 기능을 추가하면 각 사용자의 최근 검색 기록을 저장할 수 있다. 이를 분석하여 강의 검색 키워드를 맞춤형 추천 키워드로 제공할 수 있다.

9. 결론

모델 성능 평가 결과, 정확도 **0.79**, 정밀도 **1.00**, 재현율 **0.79**, **F1-Score 0.88**의 성능을 보였다. 이는 질문에 대해 관련된 강의를 정확하게 찾아내지만, 1) 복합명사를 단일 명사로 분리하여 검색하는 문제와 2) 키워드에 오타가 있거나 자모음으로 대체된 경우 이를 적절히 수정하지 못하는 한계가 있음을 시사한다. 응답 시간은 평균 **0.17초**로, 고객이 높은 만족도를 보이는 평균 반응 속도인 **2~5초**⁶보다 충분히 짧으므로 빠른 속도로 응답이 가능하다고 평가할 수 있다.

결론적으로, 해당 모델은 수업 내용에 대해 빠르고 정확한 안내를 제공하지만, 성능 개선의 필요가 있다. 모델의 정확도와 재현율이 향상된다면, 다양한 학습자의 요청을 더 정확하게 처리할 수 있어 교육 분야 서비스에 효과적으로 활용될 것으로 기대된다.

해당 모델이 교육 서비스에 응용될 수 있는 예시는 다음과 같다.

1. 개인 맞춤형 학습 경로 제공 : 학생의 질문 패턴과 학습 성향을 분석해, 그에 맞는 맞춤형 학습 경로를 제안할 수 있다. 이를 통해 학생은 자신의 학습 수준과 목표에 맞는 강의를 추천받을 수 있다.
2. 교육용 챗봇 : 학생들이 학습 중에 생기는 질문을 입력하면 실시간으로 관련 강의 및 학습 자료를 찾아 제공하는 교육용 챗봇 서비스에 적용할 수 있다.

⁶ Gnewuch, U., Morana, S., Adam, M. T. P., & Maedche, A. (2022). "Opposing effects of response time in human–chatbot interaction". *Business & Information Systems Engineering*, 64, p.773-791.

10. 소스코드

- 프로젝트의 소스코드는 별첨 자료로 제공된다.

- **Web** : 챗봇 웹사이트 구성

11. 참고문헌

- Gnewuch, U., Morana, S., Adam, M. T. P., & Maedche, A. (2022). "Opposing effects of response time in human–chatbot interaction". *Business & Information Systems Engineering*, 64.

- 김혜진 (2009). 「피아제(J. Piaget) 인지발달이론의 교육적 시사점에 비추어 본 초등학교 수학교과서의 적절성 분석: 5, 6, 7차를 중심으로」. 석사학위논문, 창원대학교 대학원.