

캡스톤 디자인 I 최종결과 보고서

프로젝트 제목(국문) : AI 챗봇을 활용한 온라인 강의 웹 서비스

프로젝트 제목(영문) : Online lecture web service using AI chatbot

프로젝트 팀(원): 학번: 20171612 이름 : 박상현

프로젝트 팀(원): 학번: 20197132 이름 : 주준하

프로젝트 팀(원): 학번: 20171768 이름 : 남승완

1. 중간보고서의 검토결과 심사위원의 '수정 및 개선 의견'과 그러한 검토의견을 반영하여 개선한 부분을 명시하시오.

- 질의응답 시스템에 관한 수정의견

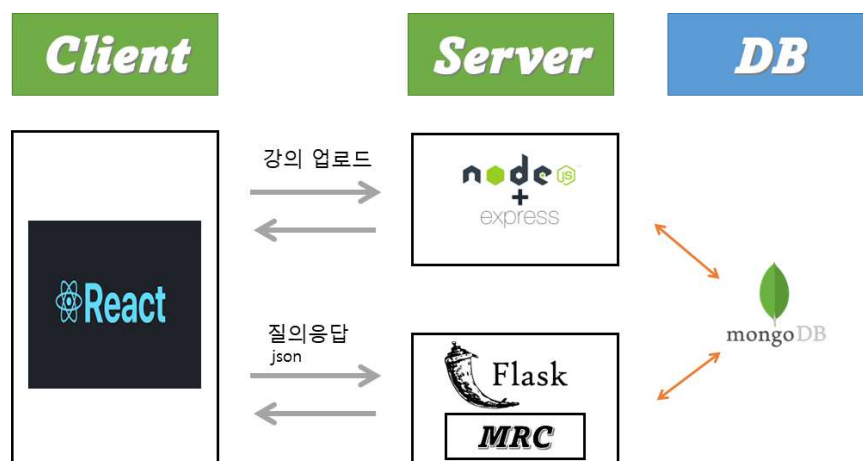
위키피디아 문서를 이용한 질의응답 챗봇은 모델의 한계, 데이터 셋의 구성 등의 이유로 실제 사용자 질문을 처리하기 힘들고 이를 해결하기에 난이도가 높다고 판단됨. 따라서 해당 기능 대신 강의자료, 계획서 등 파일에 대한 기계독해를 통해서 온라인 강의를 지원하는 새로운 기능을 제안함.

- 개선사항

KorQuad 1.0은 질문, 문서, 답변으로 구성된 데이터 셋이며 모든 답변이 단답형이라는 특성이 있다. 따라서 문장형 답변이 필요한 경우에도 단답형 답변을 출력하는 한다. 이를 개선하기 예측한 답변이 속한 문장 전체를 출력하는 방법을 사용하여 문장형 답변을 출력하도록 변경했다. 해당 방법은 답변 정확도에 대한 검증이 부족하기 때문에 다양한 테스트가 필요하며 필요시 변경하도록 다른 연구를 병행하는 중 이다. 수정의견에 따라서 위키피디아 질의응답 기능을 제거할 계획이며 강의자료, 계획서 등 파일에 대한 기계독해 기능의 설계, 구현이 완료된 상태이다.

2. 기능, 성능 및 품질 요구사항을 충족하기 위해 본 개발 프로젝트에서 적용한 주요 알고리즘, 설계방법 등을 기술하시오.

<웹>



클라이언트와 서버두개를 운용한다. 강의 업로드 서버는 DB와 연결한다. 클라이언트는 React를 통해 웹 앱을 구축하고 강의 업로드 서버는 node.js, 챗봇 질의응답에 사용한 서버는 Flask로 구현하였다. DB는 mongoDB를 사용하였다.

강의 업로드시 Json 형식을 통해 제목, 설명, 파일 이름, 동영상 Url을 보내주며 강의 계획서 파일은 재사용 할수 있도록 로컬에 저장하였다.

챗봇 질의응답은 websocket을 통해 실시간 양방향 통신 가능하도록 구현하였고 이때도 Json형식을 통해 통신 하였다.

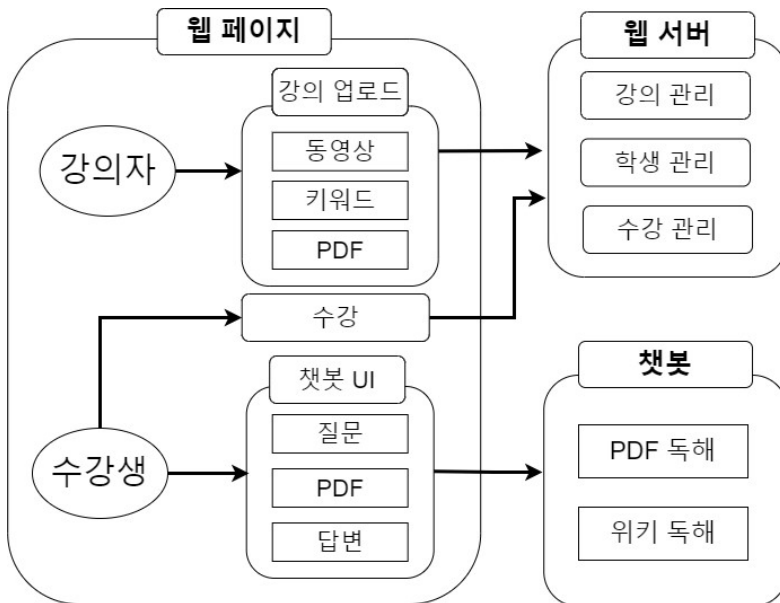
챗봇 서버는 클라이언트로부터 채팅 내용, 현재 강의 내용을 입력 받고 챗봇 모델에 입력하여 사용자에게 응답을 해준다.

<챗봇 모델>

질의응답 시스템을 구현하기 위해서 핵심 방법으로 BERT를 이용한 기계독해를 활용한다. 질문과 문단 쌍을 tokenizing해서 BERT에 입력하면 토큰에 대한 임베딩을 생성하고, 해당 토큰 임베딩을 답변의 범위를 예측하는 레이어에 입력하여 예측 결과를 출력한다. BERT는 이러한 과정을 거쳐서 질문에 대한 답변을 문단 속에서 찾아서 출력한다. 질의응답 시스템의 세부 기능 중 하나인 강의자료, 계획서 질의응답에 대한 설계는 사용자가 강의 계획서, 강의자료를 서버에 업로드하고 서버에서 업로드된 PDF 파일에서 텍스트를 추출하여 질문과 함께 BERT에 입력하고 기계독해를 수행한 뒤 예측한 답변을 사용자에게 전달하도록 구성했다. 다음은 위키 피디아 문서를 이용한 질의응답은 형태소 분석기를 이용해서 질문에서 명사, 고유명사 등을 키워드로 추출하고 해당 키워드를 위키피디아에서 검색한다. 그리고 검색된 문서를 질문과 함께 BERT에 입력하여 답변을 얻고 사용자에게 전달한다. 추가적으로 문서 로딩 시간이 답변 응답 시간으로 연결되기 때문에 검색된 문서는 로컬에 저장하여 재사용할 수 있도록 설계했다.

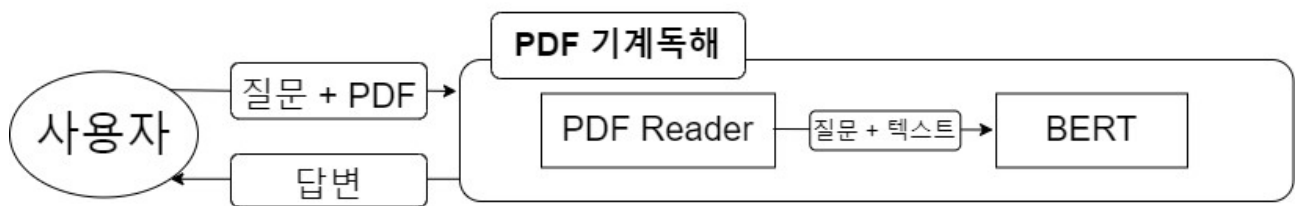
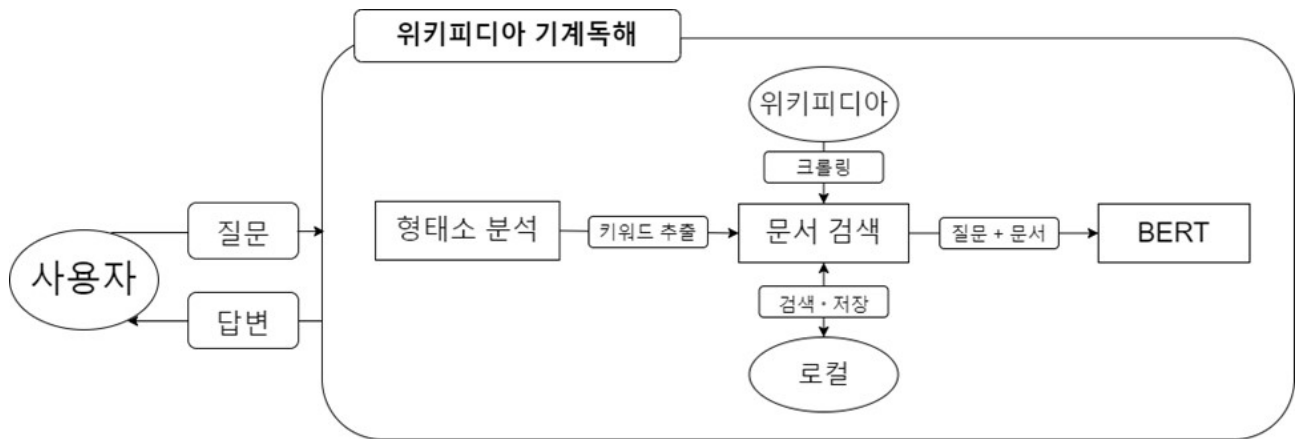
3. 요구사항 정의서에 명시된 기능 및 품질 요구사항에 대하여 최종 완료된 결과를 기술하시오.

- 전체 시스템 구성도

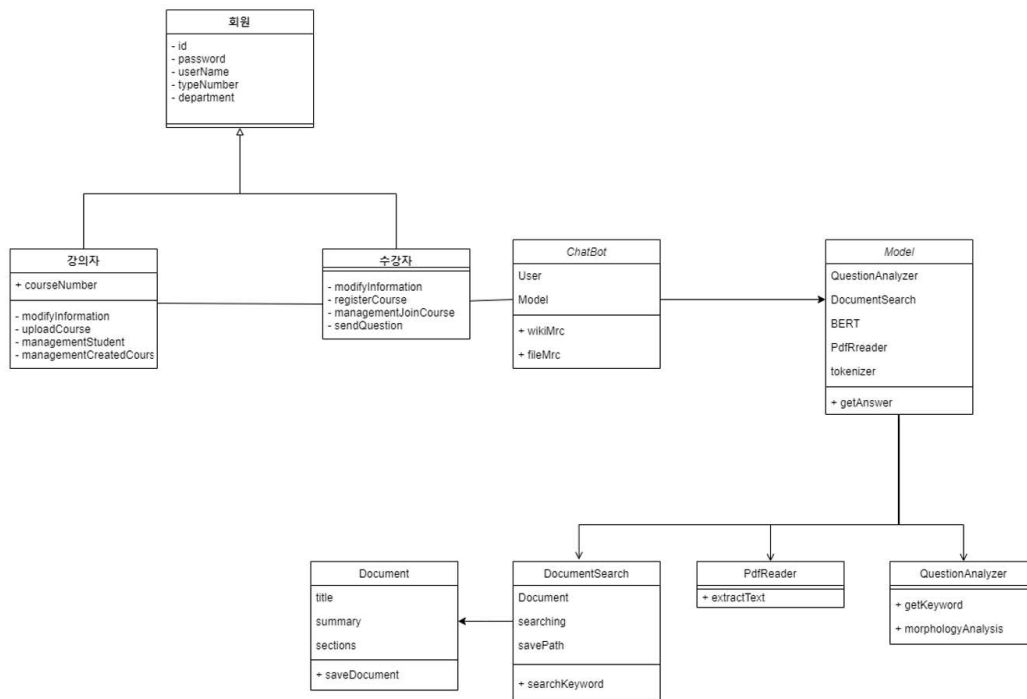


- 기능별 상세 요구사항

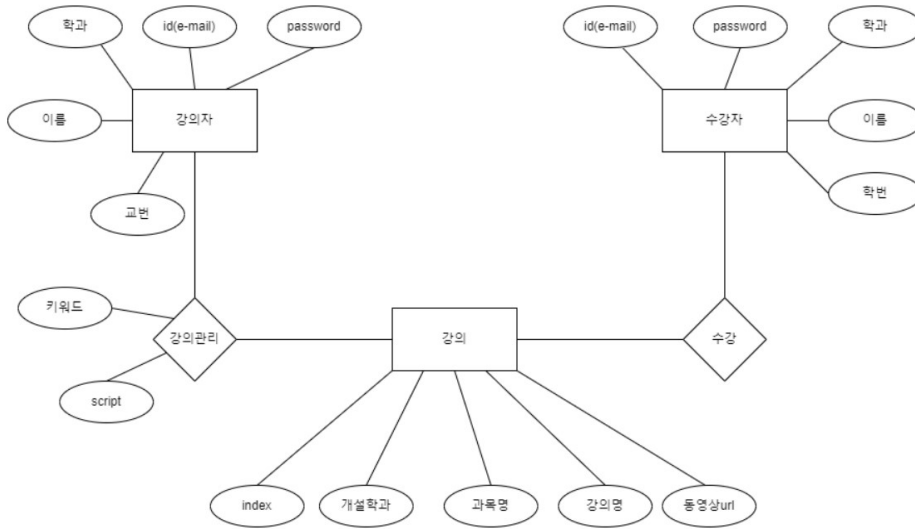
웹 기능	강의 업로드, 강좌 목록, 학습 페이지, 챗봇 페이지
PDF MRC(<u>기계독해</u>)	강의 자료(교안)를 업로드하고, <u>챗봇</u> 은 PDF의 내용을 바탕으로 <u>기계독해</u> 를 하여 수강자에게 답변을 제공함
위키피디아 MRC(<u>기계독해</u>)	사용자의 질문을 분석하고, 관련된 문서를 위키에 탐색하여 적절한 답변을 사용자에게 제공함



- 클래스 다이어그램



- E-R 다이어그램



4. 구현하지 못한 기능 요구사항이 있다면 그 이유와 해결방안을 기술하시오.

최초 요구사항	구현 여부(미구현, 수정, 삭제 등)	이유(일정부족, 프로젝트 관리미비, 팀원변동, 기술적 문제 등)
기계독해	수정	위키피디아 기계독해, PDF 기계독해로 나누어 최적의 방안을 고안
로그인 관리	미구현	일정부족
회원 및 강사 데이터 처리	미구현	일정부족
임베딩 검색기	삭제	프로젝트 범위 축소

5. 요구사항을 충족시키지 못한 성능, 품질 요구사항이 있다면 그 이유와 해결방안을 기술하시오.

분류(성능, 속도 등) 및 최초 요구사항	충족 여부(현재 측정결과 제시)	이유(일정부족, 프로젝트 관리미비, 팀원변동, 기술적 문제 등)
언어모델 성능	미흡	일정 부족.
답변 응답 시간	200 글자에 5초	기술적 문제. 위키피디아 기계독해는 wike depth 문제로 인해 찾아야 하는 문서가 많아 시간이 오래걸림. 그렇기에 PDF 기계독해에 집중하는 방향을 고안.
챗봇의 답변 정확도	EM (exact match) F1 : 80 (F1 score)	모델 학습에 대한 기술적 어려움. 다른모델 적용 고려중
강의 관리자 페이지	미흡	업로드는 정상적으로 되나 서버로 업로드 되지 않는 상황에서 사용자가 보기에 업로드 된 것처럼 보여 혼란을 줄 가능성이 있다. 예외 상황에서 사용자 UI를 고려하여 구현해야 함.

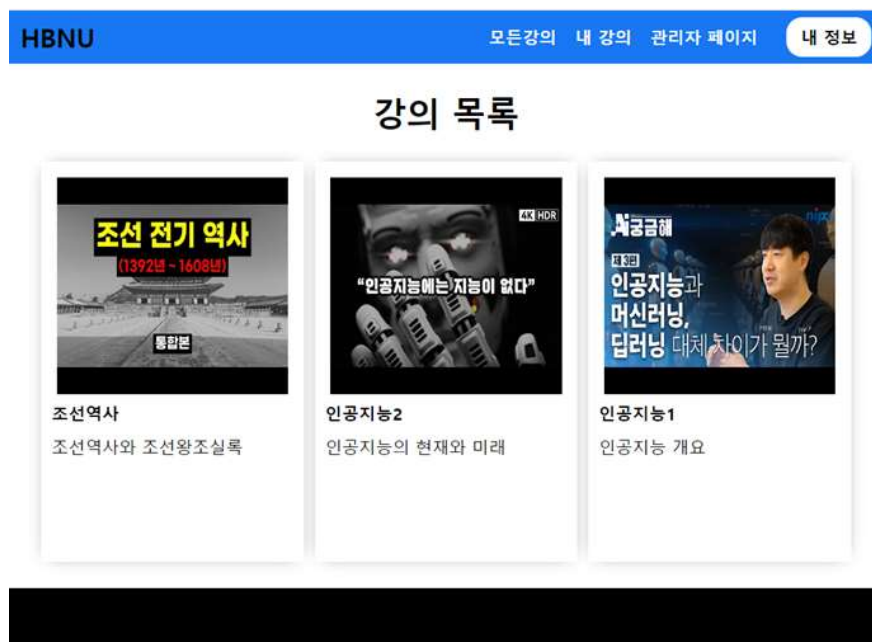
6. 최종 완성된 프로젝트 결과물(소프트웨어, 하드웨어 등)을 설치하여 사용하기 위한 사용자 매뉴얼을 작성하시오.

<상단 메뉴>



- HBNU : 페이지 로고, 클릭시 메인 페이지로 이동
- 모든강의 : 강의 서비스에 업로드된 모든 강의를 목록으로 사용자에게 보여줌
- 내 강의 : 사용자가 로그인 상태일 때 사용자가 수강하는 강의 목록을 보여줌
- 관리자 페이지 : 일반 사용자가 아닌 수강 관리자가 로그인 상태일 때 강의 업로드 페이지로 이동.
- 내 정보 : 사용자가 로그인 상태일 때 사용자 정보와, 수정을 할 수 있는 페이지로 이동.

<메인페이지>




- 메인페이지는 모든 강의 목록을 기본값 으로 보여줌.
- 강의 목록은 강의 관리자가 업로드한 강의를 보여주며 각 강의마다 이미지, 제목, 설명으로 이루어져 있음
- 강의의 이미지를 클릭하면 강의 수강 페이지로 이동.

<관리자 페이지>

HBNU 모든강의 내 강의 관리자 페이지 내 정보


강의 업로드



동영상 주소 입력하세요

제목

강의 설명

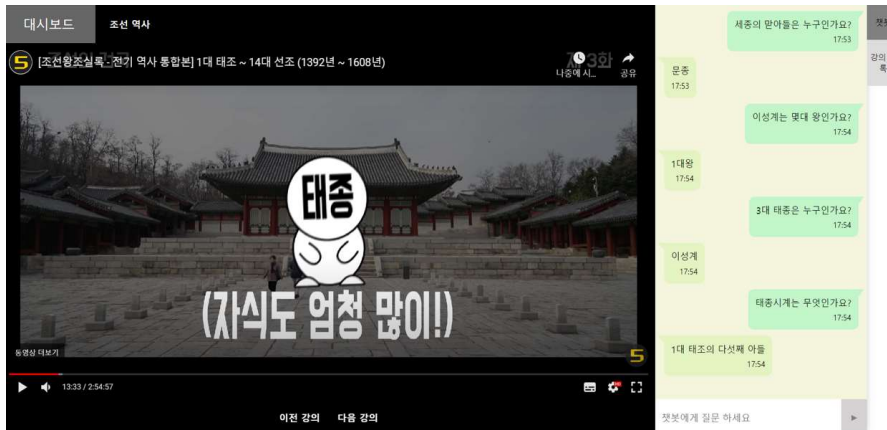


pdf 강의 계획서 업로드

업로드

- 수강 관리자가 접근가능한 관리자 페이지
- 기본적으로 강의 업로드를 할 수 있다.
- 동영상은 유튜브 동영상 주소(예시. "<https://www.youtube.com/watch?v=...>" 형식)로 업로드 가능하다. 제목과 강의설명은 text형식으로 입력 가능하다.
- pdf 강의 계획서 업로드는 pdf형식의 파일만 업로드 가능하다.
- pdf 강의 계획서 업로드는 클릭시 내 컴퓨터에 있는 파일을 업로드 할 수 있으며 파일 선택 시 title이 바뀌어 확인이 가능하다.
- 업로드는 제목, 강의 설명 둘 중 하나라도 없으면 업로드 되지 않는다.
- 동영상 주소 입력하지 않을 시 업로드 클릭하면 회색으로 표시를 해준다.
- 정상적으로 업로드가 되면 홈으로 돌아갈 수 있는 버튼으로 바뀌어 이 버튼을 클릭하면 메인 페이지로 이동한다.

<강의 수강 페이지>



- 상단 메뉴 : 메인페이지로 돌아갈 수 있는 대시보드, 강의 제목으로 이뤄져 있다.
- 하단 메뉴 : 이전강의와 다음강의로 이동할 수 있다.
- 측면 메뉴 : 챗봇과 강의 목록 둘 중 하나로 이용이 가능하다.
- 강의 영상은 강의 업로드때 사용한 유튜브 동영상 주소를 통해 유튜브 동영상으로 재생 한다.
- 챗봇은 흔히 사용하는 채팅과 같은 ui와 구조로 사용하였다.
- 챗봇 채팅은 채팅창에 값을 있어야지 입력이 활성화 된다.

7. 캡스톤디자인 결과의 활용방안

- 온라인 강의 환경은 과거에서부터 점점 증가해왔지만, 이번 코로나19 사태를 겪으면서 그 시장이 큰 규모로 커질 발판이 됐다. 불가피한 온라인 강의 환경에서 강의자와 사용자 사이에 의사소통이 제한됨으로써 강의자가 전달하고자 하는 바를 사용자가 잘 이해하지 못하는 경우가 다분하다.

AI 챗봇을 활용한 온라인 강의 웹 서비스에서 질문에 대한 답변 데이터는 검증된 많은 양의 위키피디아 데이터 오픈소스와 강의자 개인의 교안, 강의계획서 등 강의자가 전하고자 하는 자료 pdf로부터 참조되기 때문에 사용자는 챗봇에게 질문을 함으로써 올바른 정보를 답변받을 수 있다. 또한, 이 점은 서비스를 이용할 어떤 환경으로든 구축할 수 있는 확장성을 가지고 있다. 점차 늘어나는 온라인 강의 환경에서 사용자는 이 서비스를 이용해 비대면 상에서 강의자가 전하는 바를 보다 효율적으로 전달받을 수 있는 학습에 유연함을 제공받을 수 있다. 사용자는 자신이 이해하지 못하는 부분을 영상을 되감아 다시 보거나, 추가적인 자료를 찾아