

특집논문 (Special Paper)

방송공학회논문지 제29권 제5호, 2024년 9월 (JBE Vol.29, No.5, September 2024)

<https://doi.org/10.5909/JBE.2024.29.5.581>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

검색 증강 생성 기반 거대언어모델을 이용한 대학 학습관리시스템 학생 질의응답 챗봇

이재승^{a)}, 이종민^{b)}, 유제혁^{c)‡}

A Student Service Chatbot for Learning Management System at University using Retrieval-Augmented Generation-based Large Language Model

Jaeseung Lee^{a)}, Chongmin Lee^{b)}, and Jehyeok Rew^{c)‡}

요 약

학습관리시스템(LMS)은 학업의 전반적인 과정을 관리하고 지원하는 시스템으로서 대학 교육 운영에 있어 중요한 역할을 하고 있다. 이때, 대학은 LMS에서 학습 활동을 지원하기 위해 챗봇을 운영하고 있다. 하지만, 기존의 챗봇은 상담원이 질문에 대해 수기로 답변하였기에 비효율성이 발생하였다. 최근 자연어 처리와 거대언어모델(LLM)의 발전으로, 인공지능 기반 챗봇의 연구가 활발하다. 본 논문은 LMS에서의 학생 질의에 대하여 검색 증강 기반 LLM을 이용하여 자동으로 답변을 제공하는 챗봇을 제안한다. 먼저, 챗봇을 운영하며 축적된 채팅 데이터에 대해 임베딩 모델을 이용하여 검색 데이터셋을 구축한다. 이후, 새로운 질의가 챗봇에 입력되면 검색 데이터셋에서 가장 연관성이 높은 데이터를 추출하고, 이에 대한 신뢰성 검증 결과를 LLM의 입력으로 이용하여 답변을 생성한다. 본 챗봇을 다양한 질문으로 실험한 결과, 챗봇이 질문의 의도를 이해하며 명확한 답변을 제공할 수 있음을 확인하였다.

Abstract

Learning management system (LMS) plays a crucial role in managing and supporting students' educational activities. In this context, the university operates a chatbot within the LMS to aid learning activities. However, existing chatbot is less efficient because it relies on human counselors to manually respond to inquiries. With the recent development of natural language processing and large language model (LLM), there is a growing research in artificial intelligence-based chatbot. In this paper, we propose a chatbot that automatically responds to student's question in the LMS using retrieval-augmented LLM. First, we construct a retrieval dataset from the accumulated chat data using an embedding model while operating the chatbot. Then, when a new question is entered, the chatbot extracts the most relevant data from the dataset. It generates the response by inputting the confidence check result of the extracted data into the LLM. Testing the chatbot with various questions confirmed that it understands the questions' intent and provides clear answers.

Keyword : Learning management system, Large language model, Chatbot, Retrieval-augmented generation, Pre-trained language model

Copyright © 2024 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

학습관리시스템(Learning Management System, LMS)이란 학습자의 교육 및 학업의 전반적인 과정을 관리하고 지원하기 위한 소프트웨어 시스템을 말한다^[1]. 코로나19 이후 대면 수업과 온라인 강좌를 결합한 형식인 블렌디드 러닝(Blended Learning)과 전면 온라인 강좌에 대한 수요가 증가함에 따라, 대학 교육에 있어 LMS의 역할과 활용도가 크게 높아졌다^[2]. LMS는 강의 자료 공유, 과제, 평가, 토론실, 교수자와의 의사소통 도구 등의 기능을 제공하기 때문에, 학습자의 학습 참여를 촉진시키고 동시에 학습자의 학업 성취도를 높일 수 있다. 또한, 교수는 LMS를 통해 학습자로부터 강의에 대한 질문이나 피드백을 받을 수 있고, 이를 통해 강의에서 다루는 교육 콘텐츠의 내용을 보완할 수 있다. 이에, 많은 대학들이 교과 과정 운영에 있어 LMS를 도입하고 이를 적극적으로 활용하고 있다^[2].

학습자가 학업에 임하는 데에 있어 LMS를 효과적으로 활용하는 것을 장려하기 위해서는 LMS의 이용 방법을 파악할 수 있도록 하는 것이 중요하다^[3]. 예를 들어, LMS 내 토론실에서의 게시물 작성, 댓글 입력, 투표, 파일 첨부 등 다양한 기능들을 숙지한다면 LMS에서의 학습자 간의 원활한 토론 활동을 유도할 수 있다. 또한, 과제 재제출, LMS 내 온라인 강의 링크 접속, 이메일을 통한 강의 공지 등은 LMS에서 학습자가 적극적으로 활용하는 기능이기 때문에, 학습자가 본인의 학습 환경에 맞추어 해당 기능들을 적절히 설정할 수 있도록 하는 것이 중요하다. 이러한 점을 고려하여, 대학에서는 학생의 LMS 이용 방법에 대한 질의에 응답함으로써 학생의 학습 활동을 지원하는 챗봇(Chatbot)을 운영하고 있다^[4].

챗봇은 대화한다는 뜻의 채터(Chatter)와 기계인 로봇(Robot)의 합성어로, 사용자가 소프트웨어 시스템을 효율적으로 탐색하고 활용할 수 있도록 지원하는 지능형 가상 비서를 의미한다^[5]. 기존의 학생 질의응답 챗봇은 대학 기관에 소속되어 있는 근로장학생이 학생들의 질문에 대해 일일이 수기로 답변하는 형식으로 운영되었다^[6]. 하지만, 이러한 방식은 응대 담당자가 직접 매 질문의 내용을 확인하고 답변의 근거로 적합한 자료를 찾아야 하기에 효율이 떨어지는 문제가 발생한다. 예를 들어, LMS에서의 과제 제출 방법에 대한 질의가 들어오게 된다면, LMS 이용 매뉴얼에서 해당 내용을 찾고 이를 기반으로 답변을 구성하여 응답으로 제시해야 한다. 이는 대학에서의 챗봇 운영에 있어 인간의 노동력을 일정 부분 필요로 한다는 한계가 발생한다.

최근, 자연어 처리 기술(Natural Language Processing, NLP)과 거대언어모델(Large Language Model, LLM)의 발전에 따라 자동화된 챗봇 운영이 많은 주목을 받고 있다^[7]. 방대한 양의 텍스트 데이터를 사전 학습한 모델인 LLM은 주어진 텍스트의 문맥적 의미를 파악하고 이와 관련된 정보를 제공하는 데 효과적이라는 특성이 있다^[8]. 자동화된 챗봇을 통해 챗봇 구동에 있어 필요로 하는 인간의 노동력을 최소화한다면, 챗봇 운영의 효율성을 높일 수 있다. 또한, 언어 모델을 활용한 실시간 정보 검색을 통해 답변의 정확도를 향상시키는 기술인 검색 증강 생성(Retrieval-Augmented Generation, RAG)은 인공지능 기반의 자동화된 챗봇 개발을 촉진하였다^[9]. 이 기술은 사용자의 입력을 받아 관련 정보를 데이터베이스에서 검색하고, 이를 바탕으로 자연스러운 언어 응답을 생성한다. 이에 더해, 검색된 정보를 언어 모델에 통합하여, 모델이 학습 데이터의 한계를 넘어서서 실시간으로 상황에 적합한 답변을 제공할 수 있도록 한다. 이러한 RAG의 특성은 거대언어모델의 일반화 능력을 강화함으로써 모델의 응답 품질을 향상시키는 데 기여할 수 있다.

본 논문은 검색 증강 기반의 거대언어모델을 이용한 대학 LMS 질의응답 챗봇 모델 개발 기법을 제안한다. 먼저, 대학에서 챗봇 모델을 운영하며 축적된 채팅 데이터를 질의응답 형식으로 전처리하고, 가중치가 동결된(Frozen) 사전 학습 언어모델(Pre-trained Language Model, PTM)을 이

a) 고려대학교 전기전자공학부(Department of Electrical and Computer Engineering, Korea University)

b) 하버드대학교 생물통계학과(Department of Biostatistics, Harvard University)

c) 덕성여자대학교 데이터사이언스학과(Department of Data Science, Duksung Women's University)

‡ Corresponding Author : 유제혁(Jehyeok Rew)

E-mail: jhrew@duksung.ac.kr

Tel: +82-2-901-8757

ORCID: <https://orcid.org/0000-0003-3952-7582>

· Manuscript July 25, 2024; Revised August 7, 2024; Accepted August 8, 2024.

용하여 질의에 대한 임베딩(Embedding)을 도출함으로써 검색 데이터셋을 구축한다. 이후, 새로운 질의가 챗봇 모델에 입력되면, 마찬가지로 가중치가 동결된 PTM을 이용하여 임베딩을 도출하고, 이를 검색 데이터셋에서의 정보 검색에 활용하여 주어진 질의와 가장 연관성이 높은 데이터 5개를 검색한다. 정보 검색 결과, 주어진 질의와 검색 데이터셋에서 도출된 데이터 간의 유사도가 일정 기준 이상이라면 LLM이 해당 검색 데이터를 활용하여 답변을 생성하고, 유사도가 일정 기준 미만이라면 주어진 질의에 대한 답변을 제공할 수 없다는 내용의 응답을 제시한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구를 설명하고, 3장에서는 제안 기법에서 이용되는 기술과 기법의 구조에 대하여 기술한다. 4장에서는 실험에서 사용한 데이터셋을 설명하고, 제안 기법의 실험 결과를 보이며 실험 결과를 분석한다. 5장에서는 결론을 서술하며 본 논문의 끝을 맺는다.

II. 관련 연구

정확한 답변을 제공하는 챗봇 모델을 개발하기 위해, 자연어 처리 알고리즘 기반의 다양한 연구들이 제안되었다^[10-12]. Ha and Nam (2022)은 웹 문서에 기재되어 있는 공간 텍스트 정보를 이용하여 해당 위치에 웹 문서를 직접 맵핑할 수 있는 알고리즘을 제시하였다^[10]. 이를 통해, 공간 영역 및 키워드 질의 서비스를 구축함으로써 특정 위치에 대한 질의에 명확한 답변을 제공할 수 있도록 하였다. Kim and Bu (2007)는 질의에서 요구하는 정답 유형과 개념적으로 비슷한 문장을 이용하여 정답을 추출함으로써 답변의 정확도를 향상시키는 알고리즘을 제안하였다^[11]. Min and Yoo (2022)는 한국어 형태소 분석기와 Doc2Vec 기반의 문서 분류 모델을 이용하여 중학교 과학 질의응답 챗봇 모델을 개발하였다^[12]. 이후, 실제 중학교에서의 챗봇 서비스 운영을 통해 챗봇에 대한 학생들의 이용 양상을 분석하였다.

최근, LLM을 이용하여 챗봇을 개발하는 시도가 등장하였다^[13-16]. Kim and Yu (2023)는 주어진 질의에 대한 문서 검색 결과를 기반으로 한 LLM을 이용하여 한국원자력연구원의 규정을 질의응답하는 시스템을 개발하였다^[13]. 내부

규정과 관련된 다양한 질문을 통해 챗봇의 성능을 평가한 결과, 개발된 챗봇이 질문의 의도를 명확히 이해함으로써 복잡한 규정에 대해서도 사용자에게 정확하게 답변을 제공할 수 있음을 확인하였다. Kim et al. (2023)은 LLM 기반 심리상담 챗봇을 개발한 후 실제 사용자에 대한 실험을 통해 챗봇 서비스에 대한 만족도 및 효과와의 관계를 분석하였다^[14]. Yang et al. (2024)는 재료과학 분야를 다루는 공개 데이터셋인 S2ORC 데이터셋을 이용하여 메타에서 개발한 LLM인 LLaMa-2를 미세조정하였다^[15]. 이후, 실험을 통해 챗봇이 재료과학 관련 질의에 대해 명확한 답변을 제공하고 있음을 확인하였다. Birkun and Gautam (2024)은 LLM을 이용하여 심장마비 환자에 대한 응급처치 방법을 제공하는 챗봇을 개발하였다^[16].

하지만, 아직까지 대학 LMS에서의 이용 방법에 대한 질의응답을 수행하는 챗봇에 대한 연구는 활발히 수행되지 않았다. 이에, 본 논문에서는 LMS의 이용 방법에 관하여 주어진 질의와 가장 연관성이 높은 데이터를 검색하고, 이를 기반으로 한 LLM이 주어진 질의에 적합한 응답을 제시하는 방법을 제안한다.

III. 제안하는 기법

본 단원은 제안하는 기법의 전반적인 구조를 기술한다. 그림 1은 제안하는 기법의 전체 구조도를 나타낸다.

1. 텍스트 전처리

주어진 질의와 가장 유사도가 높은 정보를 선택할 수 있도록 하는 검색 데이터셋(Retrieval Dataset)을 구축하기 위하여, 채팅 데이터와 공식 Frequently Asked Questions (FAQ) 데이터를 질의응답 데이터 형식(Question-Answering Dataset)으로 전처리한다. 텍스트 전처리란, 주어진 문제의 목적에 맞추어 텍스트 데이터를 사전에 처리하는 작업을 말하며 여기에는 불필요한 문자 및 불용어 제거, 어간 및 표제어 추출 등이 포함된다^[17]. 이러한 과정은 텍스트 데이터에서의 일관성을 확보함으로써 모델이 데이터를 효과적으로 이해하고 처리할 수 있도록 한다.

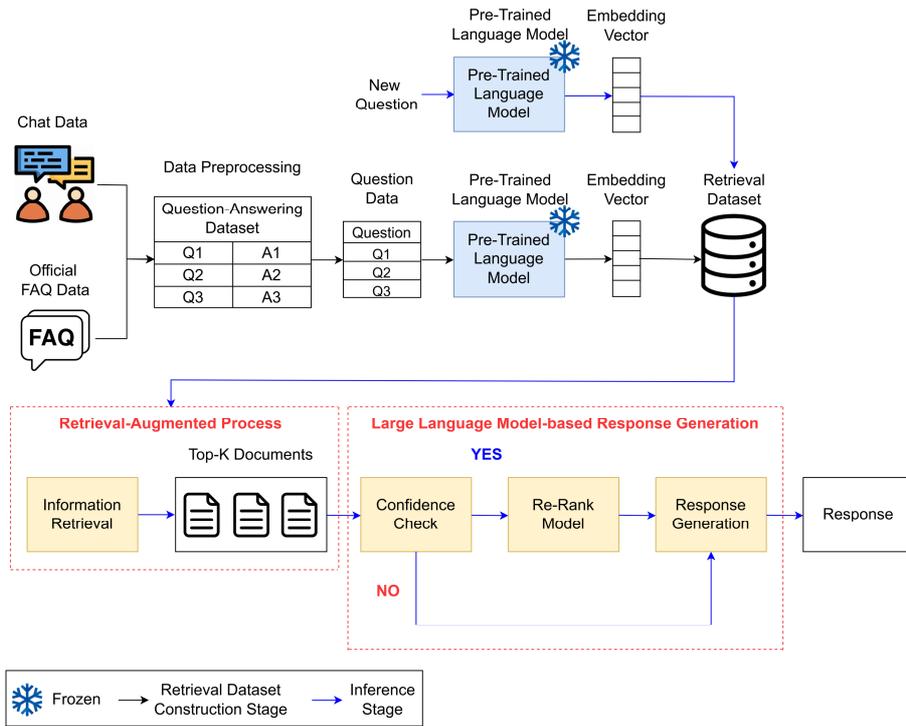


그림 1. 제안하는 기법의 전체 구조도
Fig. 1. Overview of the proposed scheme

제안하는 기법에서는 기존에 대학 LMS에서 챗봇을 운영하며 축적한 채팅 데이터를 이용한다. 따라서, LMS에서 공식적으로 제공하는 FAQ 데이터와 함께 과거 채팅 내역 데이터를 전처리한다. 이때, 공적인 내용에 대하여 범용적으로 운영할 수 있는 LMS 챗봇을 개발하기 위하여, 특정 교수자 명이나 특정 수업 이름이 언급된 채팅 내역은 제외한다. 학습자가 직접 챗봇에 질의를 제시한 특성상, 질문 데이터의 문장을 시작하는 데 있어 ‘안녕하세요’나 문장을 끝내는 데 있어 ‘감사합니다’ 등 인사말을 언급한 경우가 있다. 텍스트 전처리를 통해 구축되는 검색 데이터셋에서의 검색 정확도를 향상시키기 위해, 질문 및 응답 데이터에서의 인사말을 모두 제거한다. 또한, 한국어 데이터에만 초점을 두기 위하여, 영어로 상담이 이루어진 데이터는 제외한다. 이와 같은 기준을 준수하며 챗봇 운영에 있어 범용적으로 받아들여질 수 있는 데이터를 선택하고, 이 데이터를 질의응답 형식으로 전처리함으로써 검색 데이터셋 구축을 위한 기반 데이터를 마련한다.

2. 의미론적 임베딩 도출을 통한 검색 데이터셋 구축

3.1절에서 전처리한 데이터에 대하여 의미론적 임베딩을 도출하여 검색 데이터셋을 구축한다. 챗봇은 사용자의 질의를 입력으로 받기 때문에, 전처리 데이터에서의 질문 데이터에 대해서만 PTM의 가중치를 동결시켜 임베딩을 도출한다.

PTM은 대규모 텍스트 데이터에서 사전 학습된 언어 모델이기 때문에 주어진 텍스트 데이터의 문맥적 의미를 파악한 특성인 임베딩을 도출할 수 있다^[18]. 하지만, PTM은 규모가 큰 모델이기 때문에 새로운 데이터를 이용하여 전체 모델을 재학습시키기 위해서는 많은 양의 컴퓨터 계산 자원을 필요로 한다. 여기서, PTM의 가중치를 동결시키면서 특정 부분만 수정한다면 더 적은 자원으로도 효율적으로 주어진 작업을 수행할 수 있다. 또한, PTM은 불충분한 양의 새로운 데이터를 학습할 때 과거에 학습한 데이터의

정보를 망각하는 치명적 망각(Catastrophic Forgetting) 현상이 발생할 수 있다^[19]. 이를 방지하기 위해, 제안하는 기법에서는 PTM의 가중치를 동결함으로써 계산 자원을 절약하면서도 데이터에 대한 독립적인 표현 학습 능력을 유지한다.

채팅에서의 질문 데이터는 주로 문장으로 구성된다는 특성이 있다. 이를 고려하여, 제안 기법에서는 Sentence-BERT (SBERT)를 이용하여 주어진 문장 수준에서의 채팅 데이터로부터 의미론적인 임베딩을 도출한다^[20]. SBERT는 문장의 유사성을 측정하기 위해 BERT를 변형하여 설계한 언어 모델로, 주어진 문장 데이터에 대하여 고정된 크기의 임베딩을 도출하여 문장에 내재된 문맥적 의미를 효과적으로 파악한다는 특성이 있다. 특히, 본 논문에서는 한국어 채팅 데이터를 다루기에, SBERT 중에서도 한국어 문장 데이터를 학습한 모델인 ko-sbert-sts 모델을 이용한다^[21]. ko-sbert-sts는 기계학습 모델로 하여금 한국어 데이터의 문맥을 파악하도록 하기 위해 구축된 KorNLU 데이터셋을 활용하여 SBERT를 학습한 모델로서, 한국어 문장에 대한 임베딩을 도출하는 데 있어 활발히 이용되고 있다.

3. 정보 검색

3.2절에서 구축한 검색 데이터셋은 챗봇에 입력된 새로운 질문에 대하여 정보 검색을 수행하는 데 있어서의 기반 자료가 된다. 챗봇에 새로운 질문이 입력되면, 이 텍스트 데이터 역시 그림 2와 같은 과정을 통해 임베딩을 도출하게 된다. 이러한 임베딩에 대하여 검색 데이터셋으로부터 가장 연관성이 높은 데이터 k개를 도출한다. 새로운 질문과 가장 연관성이 높은 데이터를 탐색하기 위해 코사인 유사도(Cosine Similarity)를 이용하며, 식 (1)에 따라 주어진 질문과 각 검색 데이터 간의 유사도를 계산한다. A 는 주어진 질문에 대한 임베딩 벡터를, B 는 검색 데이터셋에서의 각 데이터에 대한 임베딩 벡터를 의미한다. n 은 임베딩 벡터의 길이이다. 코사인 유사도는 두 벡터의 내적을 벡터의 크기로 나누어 계산하기 때문에, 계산이 간단하고 속도가 빨라 대규모 데이터셋에서도 효율적으로 동작할 수 있다는 장점이 있다. 이를 통해 도출된 주어진 질문과 가장 유사도가 높은 데이터 k개는 이후 단계를 수행하는 데 있어서의

참고 자료로 활용된다.

$$\begin{aligned} \text{similarity} &= \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \\ &= \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}} \end{aligned} \quad (1)$$

4. 신뢰도 검증

주어진 질문에 대하여 3.3절을 통해 도출한 가장 유사한 데이터 k개가 질문에 대하여 적절한 내용의 답변을 제공할 수 있을 만큼 유의미한 정보를 다루고 있는지 판단한다. 만일 주어진 질문에 대하여 유사도가 낮은 데이터가 참고 자료로 제공된다면, 챗봇의 답변 제공 능력이 오히려 악화할 수 있다^[22]. 이에, 주어진 질문에 대한 유의미한 정보를 제공할 수 없는 경우, 참고 자료를 제공하지 않게 한다.

3.3절에서 계산한 주어진 질문과 검색 데이터 간의 코사인 유사도를 신뢰도(Confidence) 검증의 기준값으로 활용한다. 만일 3.3절을 통해 도출된 가장 유사한 데이터 k개 중, 첫 번째로 유사한 데이터의 코사인 유사도가 특정 기준값을 넘지 않는다면 주어진 질문에 대하여 참고 자료를 제공하지 않고 3.6절로 넘어가게 한다. 반대로, 신뢰도 검증 결과 특정 기준값 이상의 수치를 보여 검색 데이터가 유의미한 정보를 제공할 수 있다고 판단되면, 참고 자료를 제공함으로써 주어진 질문에 대한 적절한 답변을 제시할 수 있도록 한다.

5. 재순위화 모델

코사인 유사도는 텍스트 데이터에 대한 임베딩 벡터 간의 수치적 유사도를 기반으로 정보를 검색하기 때문에 이를 통해 도출된 참고 자료가 답변 제공에 있어 직접적으로 도움이 되는 단서를 제공하지 못할 수도 있다. 이에, 표 1의 프롬프트와 같이 주어진 질문과 이와 가장 유사도가 높은 k개의 데이터를 LLM에 입력하여 참고 자료를 재순위화한다. 본 단계에서는 방대한 양의 텍스트 데이터를 사전 학습하여 텍스트 데이터의 문맥적 의미를 파악하는 데 효과적

인 성능을 보인 OpenAI의 GPT-3.5의 API를 이용한다^[23]. 여러 참고 자료들을 전체적으로 평가하여 주어진 질문과의 유사도에 대한 순위를 매기는 방식인 Listwise Method를 이용하여 참고 자료를 재순위화하고^[24,25], 이를 통해 도출된 참고 자료의 유사도 순위 순서대로 3.6절의 답변 생성 프롬프트에 제공한다.

6. 답변 생성

답변 생성의 경우, 3.4절에서의 신뢰도 검증 결과에 따라

주어진 질문과 관련된 참고 자료가 유의미한 정보를 제공할 수 있는지의 여부에 따라 각기 다른 프롬프트를 LLM에 입력하여 수행한다. 답변 생성 단계 역시 3.5절과 같이 GPT-3.5의 API를 LLM으로 이용한다.

표 2는 주어진 질문에 대한 참고 자료가 유의미한 정보를 제공할 수 있을 경우에 대한 프롬프트를 나타낸다. 주어진 질문과 이와 관련된 참고 자료, 그리고 답변 작성 지시사항을 통합한다. LLM은 검색 데이터셋에서 도출된 참고 자료에 근거하여 주어진 질문에 대한 적절한 답변을 제공할 수 있도록 한다. 이는 RAG 기법을 적용한 방법론으로서 주어진 질문에 대한 참고 자료를 검토함으로써 답변을 생성하

표 1. 재순위화 모델에서의 프롬프트
Table 1. Prompt for re-rank model

Prompt Example
<p>You are a human expert on the university's Learning Management System (LMS) policy. Our main goal is to provide a precise response to the student's question about the LMS. Before doing so, we will re-rank the relevant information related to the student's question. The following are pieces of information related to the student's question:</p> <p>(1) {answer_suggestion[0]} (2) {answer_suggestion[1]} (3) {answer_suggestion[2]} (4) {answer_suggestion[3]} (5) {answer_suggestion[4]}</p> <p>Please rank this information based on its relevance to the student's question.</p>

표 2. 답변 생성에서의 프롬프트 - 사례 (1) 주어진 질의에 대한 참고 자료가 유의미한 정보를 제공할 수 있는 경우
Table 2. Prompt for response generation - Case (1) References to the given questions can provide meaningful information

Prompt Example
<p>As a chatbot within a university's Learning Management System (LMS), your role is to address students' inquiries about the LMS functionalities. To ensure high accuracy in responses, we employ the Retrieval-Augmented Generation methodology. Embedding vectors have been generated for all available Question-Answer pairs using Sentence-BERT, a notable sentence-level BERT model. This process is also applied to the queries submitted by students. By comparing the cosine similarity between the embedding vector of a student's query and those of existing Question-Answer pairs, we identify the top 5 most relevant pairs. Your input will consist of the student's question along with the 5 closest matching answers. Your objective is to synthesize this information to formulate a clear and precise response to the student's question. Note that both the input data and your responses will be in Korean.</p> <p>(1) {answer_suggestion[0]} (2) {answer_suggestion[1]} (3) {answer_suggestion[2]} (4) {answer_suggestion[3]} (5) {answer_suggestion[4]}</p> <p>Please provide your answers except for the words referring to responders such as 'student'.</p>

표 3. 답변 생성에서의 프롬프트 - 사례 (2) 주어진 질의에 대한 참고 자료가 유의미한 정보를 제공할 수 없는 경우
Table 3. Prompt for response generation - Case (2) References to the given questions cannot provide meaningful information

Prompt
<p>Please answer the questions presented as follows by saying that it is difficult to answer: {new_question}</p>

기 때문에 답변의 정확도를 향상시킬 수 있다⁹⁾.

표 3은 참고 자료가 주어진 질문에 대한 유의미한 정보를 제공할 수 없는 경우에 대한 프롬프트를 나타낸다. 주어진 질문에 대해 답변이 어렵다는 내용으로 답변을 생성하도록 한다. 이를 통해, LMS와 관련된 질문에 대해서만 적절한 답변을 제공할 수 있도록 한다.

하였다. 표 4는 본 논문에서 이용한 데이터셋의 요약을 나타내며, 표 5는 해당 데이터셋의 예시를 나타낸다. 챗봇에 축적된 44,763건의 채팅 데이터와 183건의 공식 FAQ 데이터를 전처리하여 총 2,435건의 질의응답 쌍 데이터를 이용하였다.

IV. 실험 및 결과

1. 데이터셋

본 논문에서는 2022년 1월 1일부터 2023년 6월 30일까지 실제 국내 대학에서 LMS 챗봇을 운영하며 축적된 채팅 데이터와 LMS에 기재되어 있는 공식 FAQ 데이터를 이용

2. 실험 환경

제안하는 기법의 효과성을 검증하기 위해, 주어진 질의가 챗봇에 입력되었을 때, 해당 질의에 적합한 답변이 제공되는지의 여부를 검사하는 방식으로 질의응답 챗봇의 성능을 평가하였다. 이를 위해, (1)LMS에 적합한 질문을 하였을 경우, (2)대학과 관련 있는 질문이나, LMS와는 관련 없는 질문을 하였을 경우, (3)대학과 LMS 모두와 관련 없는 질문을 하였을 경우 등 크게 3가지 유형으로 나누어 챗봇

표 4. 데이터셋 요약
 Table 4. Summary of dataset

Dataset	Period	# Samples	# Processed samples
Chat data	2022. 01. 01. - 2022. 06. 30.	24,247	729
	2022. 07. 01. - 2022. 12. 31.	10,499	693
	2023. 01. 01. - 2023. 06. 30.	10,047	830
Official FAQ data	-	183	183
Total		44,976	2,435

표 5. 데이터셋 예시
 Table 5. Example of dataset

Dataset	Type	Language	Example
Chat data	Question	Korean	과제 재제출하는 방법을 알고 싶습니다.
		English	I would like to know how to resubmit an assignment.
	Answer	Korean	교수님께서 제출 가능 횟수를 2회 이상으로 설정해놓으신 경우에 한해 재제출이 가능합니다. 새로 제출 버튼이 보이지 않는다면 교수님께서 과제제출 횟수를 1회로 설정해놓으신 것입니다. 과제 제출 횟수 설정 권한은 교수님께 있습니다. 불가피하게 과제를 다시 제출하셔야 하는 경우 교수님께 직접 연락드리셔야 할 것 같습니다.
		English	Resubmission is only possible if the professor has allowed two or more submission attempts. If you do not see the resubmit button, it means the professor has set the submission limit to one. The authority to determine the number of submission attempts lies with the professor. If you need to resubmit the assignment due to unavoidable circumstances, you will need to contact the professor directly.
Official FAQ data	Question	Korean	실시간 강의 안내를 받았는데 메뉴가 안보여요.
		English	I received a notification about a live lecture, but I can't find the menu.
	Answer	Korean	공지사항이나 이메일을 통해서 실시간 수업 강의 안내를 받으셨음에도 Collaborate, Zoom 메뉴가 보이지 않으신가요? 교수님께서 아직 메뉴를 열지 않은 상태인 것 같습니다. 교수님께 '메뉴보이기' 활성화를 요청해주세요.
		English	Did you receive a notification about the live lecture via announcement or email, but can't see the Collaborate or Zoom menu? It seems the professor hasn't opened the menu yet. Please ask the professor to enable the 'Show Menu' option."

작동 실험을 수행하였다. 3.3절에서 주어진 질의에 대하여 검색 데이터셋으로부터 가장 연관성이 높은 데이터 k개를 도출하는 것의 k는 5로 지정하였으며, 3.4절에서 신뢰도 검증의 기준값은 0.7로 설정하였다. 개발 언어는 Python 3.7.13과 PyTorch 1.13.1 프레임워크^[26]를 이용하였다.

3. 질의응답 챗봇 성능 평가

제안하는 기법의 성능을 검증하기 위해, 4.2절에서 언급한 3가지 유형 각각에 해당하는 10개의 질문을 이용하여 챗봇의 답변을 확인하였다. 생성한 텍스트와 참조 텍스트 간의 유사성을 측정하여 평가하는 지표인 **Recall-oriented understudy for gisting evaluation (ROUGE)**^[27]를 이용하여 챗봇이 제공하는 답변에 대한 성능을 정량적으로 평가하였다. 또한, 복수의 사용자들에 대한 설문조사를 통해 챗봇이 제공하는 답변에 대한 만족도를 정성적으로 평가하였다. 챗봇이 주어진 질의에 대해 항상 올바른 정답을 제공할 수 있는 것은 아니기에, 챗봇이 제공하는 답변의 양상을 분석하며 이를 보완해나가는 것이 중요하다. 그림 2는 Streamlit 1.34.0 프레임워크^[28]를 이용하여 구현한 챗봇 시스템의 웹 페이지의 동작 화면으로 사용자의 질의와 답변의 내용을

보여준다.

표 6은 LMS에 적합한 질문을 하였을 경우에 대한 챗봇의 성능을 평가한 결과이다. 과제 재제출, 청강생 등록, 이메일 변경, 수강생 명단 확인, 토론실에서의 게시물 작성 등의 방법에 대한 질의에 대하여 챗봇이 명확한 답변을 제시하고 있음을 확인할 수 있다. 특히, 답변을 생성할 때 참고한 링크가 있다면, 해당 링크도 함께 제시하고 있다. 이는 답변을 생성하는 데 있어 참고한 자료라는 객관적인 근거를 제시함으로써 제공된 답변의 신뢰도를 높일 수 있다는 장점이 있다.

표 7은 대학과 관련 있는 질문이나, LMS에는 관련 없는 질문을 하였을 경우에 대한 챗봇의 성능을 평가한 결과이다. 현재 대학의 총장, 대학에 있는 모든 단과대학의 명칭 나열, 통계학과의 한 학년 정원, 대학 입시에서의 입학 전형, 캠퍼스 내 이동 경로 등에 대한 질의에 대하여 챗봇이 주어진 질의에 대한 답변을 제공할 수 없다는 식의 답변을 제시하고 있음을 확인할 수 있다. 본 챗봇은 LMS에 특화되어 개발되었기 때문에, 대학과 관련 있다 하더라도 LMS와 관련 있지 않으면 답변을 제시하지 않도록 구성되었다. 그렇기에 LMS와 관련이 없는 질문에 대해 정상적인 답변을 제공하고 있다는 점을 확인할 수 있다.

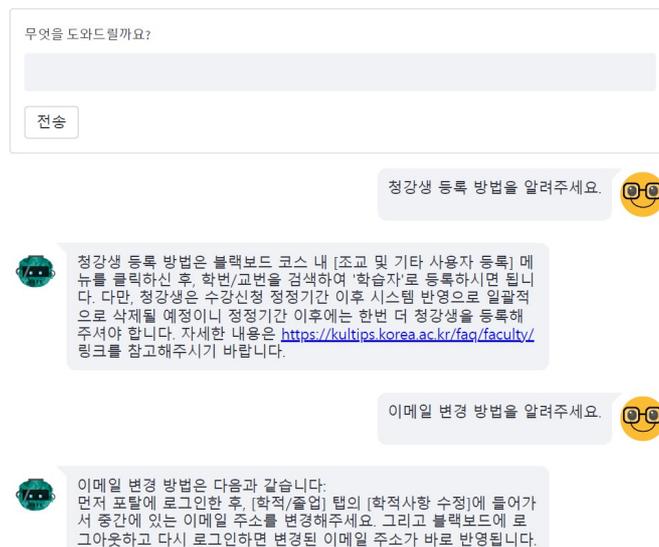


그림 2. 챗봇 시스템의 웹페이지 동작 화면
Fig. 2. Service image for the chatbot system

표 6. 제안하는 기법에서의 답변 생성 검증 - 사례 (1) LMS에 적합한 질문을 하였을 경우

Table 6. Verification of answer generation under the proposed scheme - Case (1) Asking a question suitable for the LMS

No	Type	Language	Content
1	Question	Korean	과제를 이미 제출했는데, 새롭게 제출하는 방법이 있나요?
		English	I have already submitted the assignment, but is there a way to resubmit it?
	Answer	Korean	과제를 이미 제출했을 경우, 새롭게 제출하는 방법은 없습니다. 교수님께서 제출 가능 횟수를 2회 이상 설정해놓으신 경우에 한해 재제출이 가능하며, 과제 제출 횟수 설정 권한은 교수님에게 있습니다. 불가피하게 과제를 다시 제출해야 하는 경우, 교수님께 직접 연락하시는 것이 좋습니다.
		English	If you have already submitted an assignment, there is no way to resubmit a new one. Resubmission is only possible if the professor has set the allowed number of submissions to two or more, and the authority to set the number of submissions lies with the professor. If it is necessary to resubmit the assignment, it is advisable to contact the professor directly.
2	Question	Korean	청강생 등록 방법을 알려주세요.
		English	Please tell me how to register as an auditing student.
	Answer	Korean	청강생 등록 방법은 블랙보드 코스 내 [조교 및 기타 사용자 등록] 메뉴를 클릭하신 후, 학번/교번을 검색하여 '학습자'로 등록하시면 됩니다. 다만, 청강생은 수강신청 정정기간 이후 시스템 반영으로 일괄적으로 삭제될 예정이니 정정기간 이후에는 한번 더 청강생을 등록해주셔야 합니다. 자세한 내용은 {Link}를 참고해주시기 바랍니다.
		English	To register as an auditing student in the LMS, click on the [Teaching Assistants and Other Users Registration] menu within the course. Then, search for the student ID and select 'learner' to complete the registration. Please note that auditing students will be systematically removed after the course registration correction period. Therefore, you must re-register as an auditing student once this period has passed. For more details, please refer to the provided link. {Link}
3	Question	Korean	이메일 변경 방법을 알려주세요.
		English	Please tell me how to change my email in the LMS.
	Answer	Korean	이메일 변경 방법은 다음과 같습니다. 먼저 포탈에 로그인한 후, [학적/졸업] 탭의 [학적사항 수정]에 들어가서 중간에 있는 이메일 주소를 변경해주세요. 그리고 블랙보드에 로그아웃하고 다시 로그인하면 변경된 이메일 주소가 바로 반영됩니다.
		English	The method to change your email is as follows: First, log in to the portal, then go to [Academic/Graduation] tab and enter [Modify Academic Information] to change the email address listed in the middle. Afterward, log out of the LMS and log back in to see the updated email address immediately reflected.
4	Question	Korean	블랙보드에 같이 수강하는 수강생 목록을 볼 수 있나요?
		English	Can I see a list of students enrolled in the same course on the LMS?
	Answer	Korean	블랙보드에서는 교수님 또는 조교님만이 수강생 목록을 확인할 수 있습니다. 학생분들께서는 해당 정보를 확인하실 수 없습니다. 추가적인 도움이 필요하시면 제공된 링크를 참고해주시기 바랍니다. {Link}
		English	In the LMS, only professors or teaching assistants have access to the student list. Students cannot view this information. For additional help, please refer to the provided link. {Link}
5	Question	Korean	토론실에 글을 올리는 방법을 알려주세요.
		English	Please tell me how to post in the discussion room.
	Answer	Korean	토론실에 글을 올리는 방법은 해당 과목의 블랙보드 코스에 접속한 후, 좌측에 있는 토론실을 클릭하여 작성할 수 있습니다. 만약 토론실이 비공개로 설정되어 있는 경우, 해당 과목의 교수님께 메일을 보내어 공개를 요청하시면 됩니다. 또한, 아래 링크를 참고하여 퀵 가이드나 토론 매뉴얼을 확인하시면 도움이 될 것입니다. {Link}
		English	The method to post in the discussion room involves accessing the LMS course for the specific subject and then clicking on the discussion room on the left side to write your post. If the discussion room is set to private, you can email the professor to request access. Additionally, you may refer to the quick guide or discussion manual available at the link below. {Link}

표 7. 제안하는 기법에서의 답변 생성 검증 - 사례 (2) 대학과 관련 있는 질문이나, LMS에는 관련 없는 질문을 하였을 경우

Table 7. Verification of answer generation under the proposed scheme - Case (2) Asking a question related to the university, but not related to the LMS

No	Type	Language	Content
1	Question	Korean	현재 A 대학의 총장은 누구인가요?
		English	Who is the current president of this university?
	Answer	Korean	죄송하지만, 현재 A 대학의 총장에 대한 정보는 저에게는 알 수 없는 정보입니다. 해당 정보를 확인하려면 관련 기관이나 웹사이트를 참고하시기 바랍니다.
		English	I'm sorry, but the current information about the president of the university is unknown to me. Please refer to the relevant institution or website to verify this information.
2	Question	Korean	A 대학에 있는 단과대학의 모든 명칭을 알려주세요.
		English	Please let me know the names of all the colleges in this university.
	Answer	Korean	죄송하지만, A 대학에는 다수의 단과대학이 있으며 그 명칭을 모두 알려드리기 어려운 점 양해 부탁드립니다. 더 자세한 정보를 원하신다면 A 대학 공식 홈페이지를 참고하시기 바랍니다.
		English	I apologize, but please understand that our university comprises many colleges, and listing them all would be difficult. For more information, please refer to the university's official website.
3	Question	Korean	A 대학 통계학과의 한 학년 정원은 몇 명인가요?
		English	What is the annual enrollment quota for the statistics department at this university?
	Answer	Korean	죄송하지만, A 대학 통계학과의 한 학년 정원에 대한 정보를 알고 있는 것이 아니기 때문에 정확한 답변을 드릴 수 없습니다. 해당 정보를 확인하려면 A 대학의 공식 웹사이트나 학과 사무실에 문의하시는 것이 좋을 것입니다. 부족한 도움을 드려 죄송합니다.
		English	I apologize, but I can't give you an exact answer because I don't know any information about the quota for a year in the statistics department. To check that information, you might want to contact the official university website or department office. Sorry for the lack of help.
4	Question	Korean	A 대학에 있는 입학전형에 대해 알려주세요.
		English	Please tell me about the admissions process at the university.
	Answer	Korean	죄송하지만, A 대학의 입학전형에 대한 구체적인 정보는 정확히 알려드리기 어렵습니다. A 대학의 입학전형은 매년 변동할 수 있으며, 학과나 학부에 따라 다양한 입학요건과 절차가 적용될 수 있기 때문입니다. A 대학의 공식 홈페이지나 입학처 등에서 최신 정보를 확인하시는 것을 권장드립니다.
		English	I'm sorry, but it's difficult to give you specific information about the university's admissions process. This is because the university's admissions process may vary every year, and various admissions requirements and procedures may be applied to each department. It is recommended to check the latest information on the university's official website or admissions office.
5	Question	Korean	A 대학의 정문에서 중앙도서관까지 가는 방법을 알려주세요.
		English	Please tell me how to get from the main building to the central library.
	Answer	Korean	죄송하지만, 제가 도와드릴 수 있는 정보가 아닙니다. 다른 교통안내 시스템을 이용하시거나 현지 주민에게 직접 문의하시는 것이 도움이 될 수 있습니다.
		English	I'm sorry, but this is not the information I can help you with. It may be helpful to use another traffic guidance system or contact your local residents directly.

표 8. 제안하는 기법에서의 답변 생성 검증 - 사례 (3) 대학과 LMS 모두와 관련 없는 질문을 하였을 경우

Table 8. Verification of answer generation under the proposed scheme - Case (3) Asking a question not related to both the LMS and university

No	Type	Language	Content
1	Question	Korean	현재 우리나라의 대통령은 누구인가요?
		English	Who is the president of our country?
	Answer	Korean	죄송하지만, 저는 현재 시대에 대한 정보를 알 수 없기 때문에 대통령의 정보를 제공할 수 없습니다. 다른 질문이 있으시면 도와드릴 수 있습니다.
		English	I apologize, but I am unable to provide information about the president as I do not have access to current data. If you have any other questions, I'd be happy to assist you.
2	Question	Korean	Formula 1에 참여하는 팀들의 목록을 알려주세요.
		English	Please let me know the list of teams participating in Formula 1.
	Answer	Korean	죄송하지만, Formula 1에 참여하는 팀들의 목록은 매 시즌 변동되기 때문에 정확한 목록을 제시하기 어렵습니다. 현재 시즌에 참여하는 팀들의 목록은 공식 F1 웹사이트나 관련 뉴스 등을 참고하시기 바랍니다.
		English	I'm sorry, but the list of teams participating in Formula 1 changes every season, so it's difficult to provide an accurate list. For the list of teams participating in the current season, please refer to the official F1 website or related news.
3	Question	Korean	거대언어모델이란 무엇인가요?
		English	What is the large language model?
	Answer	Korean	죄송하지만, 거대언어모델에 대한 구체적인 정보나 답변을 제공하기 어렵습니다. 더 자세한 정보를 찾아보시거나 전문가에게 문의하시는 것을 권장드립니다.
		English	Sorry, it is difficult to provide specific information or answers about the large language model. We recommend you to find more information or contact experts.
4	Question	Korean	제가 해외에서 여권을 잃어버렸습니다. 해외에서 여권을 잃어버리면 어떻게 해야 하는지 알려주세요.
		English	I lost my passport while overseas. Please tell me what I can do for a lost passport abroad.
	Answer	Korean	죄송하지만, 그 질문에 대해서는 답변드릴 수가 없습니다. 다른 질문이 있으시면 제가 도와드릴 수 있습니다.
		English	I'm sorry, but I can't answer that question. If you have any other questions, I can help you.
5	Question	Korean	Transformer에서의 Attention mechanism이 무엇인가요?
		English	What is the attention mechanism in transformer?
	Answer	Korean	죄송하지만, 그 질문에 대한 답변은 어려워져 설명드릴 수가 없습니다. 다른 질문이 있으시면 언제든지 물어보시고, 최선을 다해 도와드리겠습니다.
		English	I'm sorry, but I can't explain the answer to that question because it's difficult. If you have any other questions, feel free to ask, and I'll do my best to assist you.

표 8은 대학과 LMS 모두 관련 없는 질문을 하였을 경우에 대한 챗봇의 성능을 평가한 결과이다. LLM의 정의, 현재 Formula 1 시즌에 참여하고 있는 팀들의 명칭 나열, 해외에서 여권을 분실하였을 시의 대처 방안 등에 대한 질의에 대하여, 표 7과 마찬가지로 주어진 질의에 대한 답변을 제공할 수 없다는 답변을 제시하고 있음을 확인할 수 있다. 이를 통해, 주어진 질문의 LMS 관련 여부를 사전에 판별함으로써 답변의 정확도를 향상시킬 수 있음을 확인할 수 있다.

정량적 평가 지표인 ROUGE를 활용한 답변 성능 평가에

있어서는, 표 6과 같이 검색 데이터셋에서 참조 데이터가 존재하는 경우인 LMS에 적합한 질문을 하였을 경우에 대해 평가를 진행하였다. 그중, ROUGE-1, ROUGE-2, ROUGE-L을 평가 지표로 활용하였다^[27,29]. ROUGE-1은 단어 단위의 일치율을 통해 생성된 텍스트의 전반적인 정확도를 평가하고, ROUGE-2는 두 단어 단위의 일치율을 통해 텍스트의 유창성과 문맥적 일치를 평가한다. ROUGE-L은 최장 공통 부분 문자열(Longest common subsequence, LCS)을 기반으로 하여 문장의 순서를 고려하면서 전체적인 유사성을 평가한다. 이러한 평가 지표들은 Precision,

Recall, F1-Score를 사용하여 생성된 텍스트가 참조 텍스트와 얼마나 잘 일치하는지를 측정하며, 챗봇이 제공한 답변의 성능을 종합적으로 평가하는데 중요한 역할을 한다. 식 (2)-(4)는 각각 ROUGE-1의 Precision, Recall, F1-Score의 공식을 나타낸다. ROUGE-1의 Precision은 생성된 텍스트의 Unigram 중 참조 텍스트에 존재하는 Unigram의 비율을 나타내며, ROUGE-1의 Recall은 참조 텍스트의 Unigram 중 생성된 텍스트에 존재하는 Unigram의 비율을 나타낸다. ROUGE-1의 F1-Score는 Precision과 Recall의 조화 평균이다. ROUGE-1에서의 Unigram을 Bigram과 LCS로 변경한 것이 각각 ROUGE-2와 ROUGE-L의 공식이 된다.

$$ROUGE-1\ Precision = \frac{Number\ of\ overlapping\ unigrams}{Number\ of\ unigrams\ in\ generated\ text} \quad (2)$$

$$ROUGE-1\ Recall = \frac{Number\ of\ overlapping\ unigrams}{Number\ of\ unigrams\ in\ reference\ text} \quad (3)$$

$$ROUGE-1\ F1-Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

표 9, 10, 11은 각각 LMS에 적합한 질문을 하였을 경우

에 있어 챗봇이 생성한 답변에 대하여 ROUGE-1, ROUGE-2, ROUGE-L을 이용하여 답변의 성능을 평가한 것을 나타낸다. 제안하는 기법에서 Confidence check와 Re-rank module을 제외한 것과 제안하는 기법 사이의 답변 제공 성능을 비교하였다. 실험에 활용한 10개의 질문에 대하여 각각 평가 지표를 계산한 뒤, 이를 평균화하였다. 그 결과, 주어진 질의에 대하여 검색 데이터셋을 통해 도출한 가장 유사한 참조 데이터에 대하여 신뢰도 검증과 재순위화 모델을 거친 이후 답변을 생성한 것이 더 좋은 성능의 답변을 제공할 수 있음을 확인할 수 있었다.

본 논문에서 개발한 챗봇이 제공한 답변에 대하여 30명의 학생들을 대상으로 정성적으로 평가하였다. 표 6, 7, 8과 같이 챗봇이 제공한 답변에 대하여, 5점 리커트 척도^[30]를 이용하여 정확한 답변의 제공 여부, 자연스러운 대화 여부, 그리고 전반적인 만족도를 조사하였다. 5점 리커트 척도는 매우 불만족, 불만족, 보통, 만족, 매우 만족으로 구성되었으며, 만족과 불만족의 응답 옵션을 균형 있게 제공하기 때문에 응답자가 명확히 답변할 수 있다는 장점이 있다. 표 12는 설문조사를 통한 답변의 성능 평가 결과를 나타낸다. 정확한 답변 제공 여부, 자연스러운 대화 여부, 전반적인 만족도 등 모든 기준에 대해 매우 불만족과 불만족에 해당

표 9. ROUGE-1을 이용한 답변 성능 평가
Table 9. Evaluation of response using ROUGE-1

Method	Metric	ROUGE-1		
		Precision	Recall	F1-Score
w/o Confidence check and Re-rank module		0.3308	0.3633	0.2759
Our method		0.4326	0.4763	0.3642

표 10. ROUGE-2를 이용한 답변 성능 평가
Table 10. Evaluation of response using ROUGE-2

Method	Metric	ROUGE-2		
		Precision	Recall	F1-Score
w/o Confidence check and Re-rank module		0.2088	0.2409	0.1660
Our method		0.2983	0.3441	0.2372

표 11. ROUGE-L을 이용한 답변 성능 평가
Table 11. Evaluation of response using ROUGE-L

Method	Metric	ROUGE-L		
		Precision	Recall	F1-Score
w/o Confidence check and Re-rank module		0.2919	0.3295	0.2461
Our method		0.3880	0.4337	0.3291

표 12. 설문조사를 통한 답변 성능 평가
 Table 12. Evaluation of response through survey

Metric	Score	Very dissatisfied	Dissatisfied	Neutral	Satisfied	Very Satisfied
Accuracy of provided answers		0.00%	3.33%	26.66%	56.66%	13.33%
Naturalness of conversation		0.00%	6.66%	33.33%	50.00%	10.00%
Overall satisfaction		0.00%	3.33%	23.33%	63.33%	10.00%

하는 비율에 비해 만족과 매우 만족에 해당하는 비율이 많은 것으로 보아 대체로 챗봇이 제공하는 답변에 대해 학생들이 만족하고 있다는 점을 파악할 수 있었다.

V. 결론

본 논문에서는 검색 증강 기반 LLM을 이용한 대학 LMS 학생 질의응답 챗봇 개발 기법을 제안하였다. 제안하는 기법의 효과성을 검증하기 위하여, 3가지 유형의 질의에 따른 챗봇의 답변을 분석하였다. 실험 결과, 제안 기법은 LMS와 관련된 질의에 대해서는 적절한 답변을 제시하고 있고, LMS와 관련이 없는 질의에 대해서는 해당 질의에 대한 답변을 제공할 수 없다는 답변을 제시하고 있음을 확인하였다. 이는 주어진 질의에 대한 단순한 정보 검색을 넘어, 재순위화 모델을 통해 질의와 가장 연관성이 높은 정보를 탐색하고 이를 기반으로 답변을 생성함으로써 제공되는 답변의 정확도를 높였다고 해석할 수 있다.

인공지능 기반 챗봇이 항상 완벽하고 정확한 답변을 제공하는 것은 어렵기 때문에, 사용자의 챗봇 이용 행태를 분석함으로써 지속적으로 챗봇 시스템을 발전시켜 나가는 것이 중요하다. 즉, 사용자의 챗봇에 대한 피드백을 바탕으로 주어진 질의에 대한 챗봇의 답변 품질을 향상시켜 나가는 것이 중요하다. 따라서, 향후 연구로 사용자의 질의에 대하여 참고할 수 있는 자료를 검색 데이터셋에 추가함에 따라 더욱 정확한 답변을 제공할 수 있도록 하고자 한다. 또한, LMS뿐만 아니라 대학 내 다양한 시스템에서 학생의 질의에 대해 자동화된 답변을 제시할 수 있는 챗봇을 추가적으로 개발할 계획이다. 이러한 시스템은 데이터 기반의 스마트 대학 캠퍼스를 구축하는 데 있어 큰 역할을 할 수 있을 것으로 기대된다.

참고 문헌 (References)

- [1] H. Lee, J. Lee, and Y. Cha, "An Analysis on the Difference of LMS Activities according to Academic Achievement of College Students in Off-line Courses", *Journal of Korean Association for Educational Information and Media*, Vol.25, No.1, pp.201-222, January 2019. doi: <https://doi.org/10.15833/KAFEIAM.25.1.201>.
- [2] A. Kim, "Analysis of Learner Behavior and Learning Performance using LMS Big Data in the COVID-19: Focused on J-University", *Journal of learner-centered curriculum and instruction*, Vol.21, No.6, pp.565-579, March 2021, doi:10.22251/jlcci.2021.21.6.565.
- [3] M. Choi, "A Study on the Interactivity of Smart LMS in a University Class", *Journal of Digital Convergence*, Vol.17, No.3, pp.395-404, March 2019. doi: <https://doi.org/10.14400/JDC.2019.17.3.395>.
- [4] D. Jo, "The Needs Analysis and Dashboard Design for Building Customized Student Support Platform Based on University Learning Analysis Data", *Doctoral dissertation of Dong-a University, Busan, Korea.*, December 2021.
- [5] J. Bang, B. Lee, and P. Park, "A Study on Guidelines for AI Ethics in the Process of Aligning Conversation Topics by Conversational Bot", *Journal of Broadcast Engineering*, Vol.28, No.5, pp.564-577, September 2023. doi: <https://doi.org/10.5909/JBE.2023.28.5.564>.
- [6] S. Ahn and S. Lee, "Improving University Homepage FAQ Using Semantic Network Analysis", *Journal of Digital Convergence*, Vol.16, No.9, pp.11-20, September 2018, doi:10.14400/JDC.2018.16.9.011.
- [7] J. Bang, B. Lee, and P. Park, "Framework Development Approach for Bias Assessment of AI-based Conversational Chatbot using Large-scale Language Model", *Journal of Broadcast Engineering*, Vol.28, No.5, pp.545-563, September 2023. doi: <https://doi.org/10.5909/JBE.2023.28.5.545>.
- [8] Y. Zhu, H. Yuan, S. Wang, J. Liu, W. Liu, C. Deng, H. Chen, Z. Dou, and J. Wen, "Large Language Models for Information Retrieval: A Survey", *arXiv preprint, arXiv:2308.07107*, August 2023.
- [9] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Kuttler, M. Lewis, W. Yih, T. Rocktaschel, S. Riedel, and D. Kiela, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks", *Proceedings of the International Conference on Neural Information Processing Systems, Virtual*, pp.9459-9474, 2020.
- [10] T. Ha and K. Nam, "Techniques for Location Mapping and Querying of Geo-Texts in Web Documents", *Journal of Korea Society of Industrial Information Systems*, Vol.27, No.3, pp.1-10, June 2022,

- doi:10.9723/jksis.2022.27.3.001
- [11] H. Kim and K. Bu, "A Query Expansion Technique using Query Patterns in QA Systems", *Journal of Korea Society of Industrial Information Systems*, Vol.12, No.1, pp.1-8, March 2007.
- [12] G. Min and J. Yoo, "Development of a middle school science Q&A chatbot using Doc2Vec and analysis of student's queries", *SNU Journal of Education Research*, Vol.31, No.3, pp.115-145, September 2022.
doi: <https://doi.org/10.54346/sjer.2022.31.3.115>
- [13] H. Kim and Y. Yu, "Development of a Regulatory Q&A System for KAERI Utilizing Document Search Algorithms and Large Language Model", *Journal of Korea Society of Industrial Information Systems*, Vol.28, No.5, pp.31-39, March 2023.
doi: <https://doi.org/10.9723/jksis.2023.28.5.031>
- [14] T. Kim, M. Kim, Y. Kim, and S. Hong, "Large Language Model-Based Evaluation for Answering in Generation-Based and Retrieval-Based Chatbots", *Korea Software Congress*, Busan, Korea, pp.1523-1525, December 2020.
- [15] X. Yang, S. Wilson, and L. Petzold, "Quokka: An Open-source Large Language Model ChatBot for Material Science", *arXiv preprint*, arXiv:2401.01089, January 2024.
- [16] A. Birkum and A. Gautam, "Large Language Model-based Chatbot as a Source of Advice on First Aid in Heart Attack", *Current Problems in Cardiology*, Vol.49, No.1, 102048, January 2024.
doi: <https://doi.org/10.1016/j.cpcardiol.2023.102048>
- [17] J. Jeong, M. Jee, M. Go, H. Kim, H. Lim, Y. Lee, and W. Kim, "Related Documents Classification System by Similarity between Documents", *Journal of Broadcast Engineering*, Vol.24, No.1, pp.77-86, January 2019.
doi: <https://doi.org/10.5909/JBE.2019.24.1.77>.
- [18] X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang, "Pre-trained Models for Natural Language Processing: A Survey", *arXiv preprint*, arXiv:2003.08271, March 2020.
- [19] Y. Luo, Z. Yang, F. Meng, Y. Li, J. Zhou, and Y. Zhang, "An Empirical Study of Catastrophic Forgetting in Large Language Models During Continual Fine-tuning", *arXiv preprint*, arXiv:2308.08747, August 2023.
- [20] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks", *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Hong Kong, China, pp.3982-3992, November 2019.
- [21] J. Gan, "ko-sbert-sts", *HuggingFace*, [Available Online] <https://huggingface.co/jhgan/ko-sbert-sts> (Access on April 01, 2024)
- [22] S. Yan, J. Gu, Y. Zhu, and Z. Ling, "Corrective Retrieval Augmented Generation", *arXiv preprint*, arXiv:2401.15884, January 2024.
- [23] OpenAI, "GPT-4 Technical Report", *arXiv preprint*, arXiv:2303.08774, March 2023.
- [24] W. Sun, L. Yan, X. Ma, S. Wang, P. Ren, Z. Chen, D. Yin, and Z. Ren, "Is chatgpt good at search? investigating large language models as re-ranking agents.", *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Singapore, Singapore, pp.14918-14937, December 2023.
- [25] X. Ma, X. Zhang, R. Pradeep, and J. Lin, "Zero-shot listwise document reranking with a large language model", *arXiv preprint*, arXiv:2305.02156, May 2023.
- [26] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. Devito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An Imperative Style, High-Performance Deep Learning Library", *Proceedings of the Advances in Neural Information Processing Systems*, Vancouver, Canada, pp.8026-8037, December 2019.
- [27] C. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries", *Annual Meeting of the Association for Computational Linguistics*, Barcelona, Spain, pp. 74-81, July 2004.
- [28] M. Kuzlu, F. Catak, S. Sarp, U. Cali, and O. Gueler, "A Streamlit-based Artificial Intelligence Trust Platform for Next-Generation Wireless Networks", *arXiv preprint*, arXiv:2211.12851, November 2022.
- [29] M. Barbella and G. Tortora, "Rouge Metric Evaluation for Text Summarization Techniques", Available at SSRN: <http://dx.doi.org/10.2139/ssrn.4120317>, 2022.
- [30] A. Joshi, S. Kale, S. Chandel, and D. K. Pal, "Likert Scale: Explored and Explained", *Current Journal of Applied Science and Technology*, Vol.7, Issue 4, pp. 396-403, 2015.

저 자 소 개



이 재 승

- 2023년 2월 : 고려대학교 통계학과 학사
- 2023년 3월 ~ 현재 : 고려대학교 전기전자공학과 석사과정
- ORCID : <https://orcid.org/0009-0007-5863-9367>
- 주관심분야 : 시계열 예측, 거대언어모델, 빅데이터 분석

저 자 소 개



이 종 민

- 2024년 2월 : 고려대학교 통계학과 학사
- 2024년 8월 ~ 현재 : 하버드대학교 생물통계학과 석사과정
- ORCID : <https://orcid.org/0009-0000-0027-6515>
- 주관심분야 : 의료 인공지능, 기계학습, 거대언어모델, 인과추론



유 제 혁

- 2020년 8월 : 고려대학교 전기전자컴퓨터공학과 박사
- 2023년 2월 : ㈜기아 책임매니저
- 2023년 3월 ~ 현재 : 덕성여자대학교 데이터사이언스학과 조교수
- ORCID : <https://orcid.org/0000-0003-3952-7582>
- 주관심분야 : 빅데이터, 패턴인식, 인공지능