

교육학박사 학위논문

우리말 제목
한 줄 떼려면 중간에 백슬래시 두 번 넣으세요
INPUT YOUR TITLE IN ENGLISH

2025년 2월

서울대학교 대학원
과학교육과 물리전공
홍길동

우리말 제목
한 줄 떼려면 중간에 백슬래시 두 번 넣으세요
INPUT YOUR TITLE IN ENGLISH

지도교수 김 영 희
이 논문을 교육학박사 학위논문으로 제출함
2024년 12월

서울대학교 대학원
과학교육과 물리전공
홍 길 동

홍 길 동의 교육학박사 학위논문을 인준함
2025년 2월

위 원 장 김 영 희 (인)

부 위 원 장 김 영 희 (인)

위 원 김 영 희 (인)

위 원 김 영 희 (인)

위 원 김 영 희 (인)

국문초록

최근 LLM(대규모 언어 모델, Large Language Model)의 발전으로 교육용 챗봇 개발 가능성이 증대되었으나, 기존의 교육용 챗봇은 학교 맥락 반영이 부족하고 상시 사용이 어려운 한계가 있었다. 이에 본 연구에서는 ...

...

...

주요어: 주요어1, 주요어2, 주요어3, 주요어4, 주요어5, 주요어6

학번: 2020-30000

차례

국문 초록

차례	i
----	---

그림 차례	iii
-------	-----

표 차례	iv
------	----

1 서론	1
------	---

1.1 연구의 배경 및 필요성	1
------------------	---

1.2 연구의 목적과 문제	1
----------------	---

1.3 연구과정의 개요	2
--------------	---

1.4 용어의 정의	4
------------	---

1.4.1 용어 1	4
------------	---

1.4.2 용어 2	4
------------	---

2 이론적 배경	5
----------	---

2.1 오픈소스 한국어 지원 LLM	5
---------------------	---

2.2 RAG (Retrieval Augmented Generation)	5
------------------------------------------	---

2.2.1 환각 현상 감소를 위한 RAG 적용	5
---------------------------	---

2.2.2 여기에 subsection 제목을 쓰세요	6
------------------------------	---

3	연구 방법	8
3.1	선행연구에서 수집한 자료의 재범주화 및 비교분석	8
3.2	챗봇 개발	8
3.2.1	LLM에 RAG를 적용한 생성형 AI 챗봇 개발 사례	8
3.2.2	프로토타입 개발 과정	8
4	연구 결과	13
4.1	section 제목	13
4.1.1	subsection 제목	13
5	결론 및 논의	14
5.1	결론 및 논의	14
5.2	연구의 한계	15
5.3	향후 연구 방향	15
A	연구에서 사용했던 약어 목록	17
B	IRB 결과 통보서와 연구 참여자 모집 문건, 설문지	19
	Abstract	20

그림 차례

3.1 LLM에 RAG를 적용하여 답변을 생성하는 과정	9
3.2 교사와 학생의 평가를 반영하여 개선한 시스템 프롬프트	10

표 차례

3.1	두 가지 학교 맞춤형 과학 질문-답변 챗봇의 사용 기록 및 사용자 평가 내용 의 개요	11
3.2	LLM별 RAG 적용 실험 조건	11
3.3	학생 질문을 주제별로 분류한 결과(민경모·유준희, 2022)	12

제 1 장

서론

1.1 연구의 배경 및 필요성

최근 자연어 처리(Natural Language Processing, NLP) 기술은 트랜스포머 기반의 거대 언어 모델(Large Language Model, LLM)로 발전하며 교육 분야에서의 활용 가능성을 크게 넓혔다. 학생들의 질문은 과학 학습의 핵심 과정으로 교사의 적절한 응답은 학습 효과를 크게 높인다(Chin and Osborne, 2008; Eshach et al., 2014).

...

1.2 연구의 목적과 문제

본 연구의 목적은 ...이다. 연구 문제는 다음과 같다.

1. 연구문제 1
2. 연구문제 2

3. 연구문제 3

(a) 세부 연구문제 1

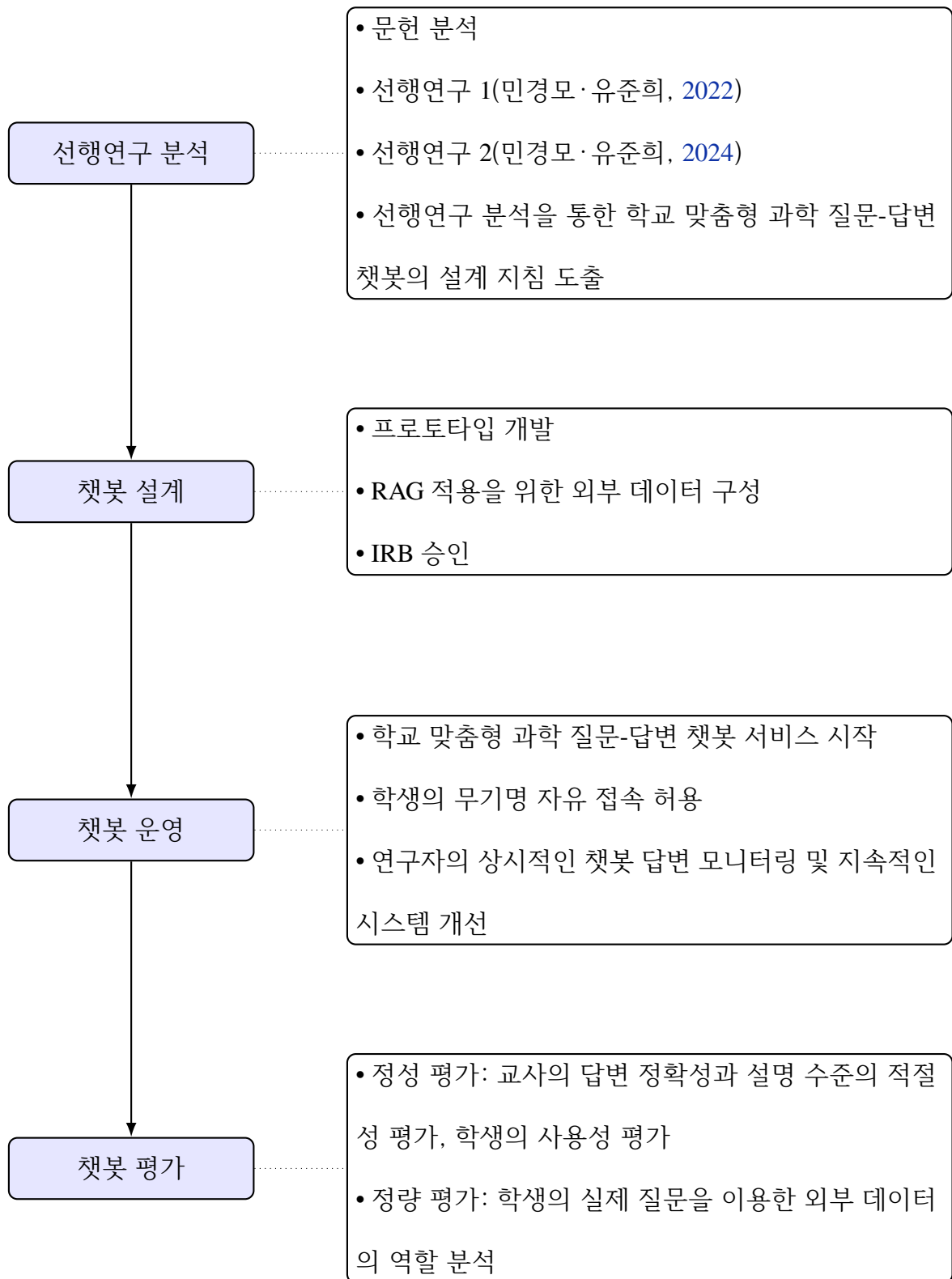
(b) 세부 연구문제 2

1.3 연구과정의 개요

본 연구는 LLM을 이용하여 학생이 학교 현장에서 실제로 사용할 수 있는 과학 질문-답변 챗봇을 개발하고 평가한 연구로서 민경모·유준희(2022)와 민경모·유준희(2024)의 후속 연구이기도 하다.

...

연구 과정의 개요는 다음과 같이 도식화하였다.



1.4 용어의 정의

1.4.1 용어 1

...

1.4.2 용어 2

...

제 2 장

이론적 배경

2.1 오픈소스 한국어 지원 LLM

각주 작성 예시

거대 언어 모델^a

...

2.2 RAG (Retrieval Augmented Generation)

2.2.1 환각 현상 감소를 위한 RAG 적용

참고문헌 삽입 예시

cite와 citep를 적절하게 사용하세요. cite는 괄호를 붙이지 않고, citep는 괄호를 붙입니다.

^a거대 언어 모델은 Large Language Model을 번역한 용어로 매우 많은 매개변수를 사용하여 깊은 신경망 (Deep Neural Network)을 구성하여 아주 큰 데이터를 사전학습 시킨 모델이다(Chang et al., 2024).

RAG는 LLM 자체는 그대로 두되 LLM이 벡터 저장소에 있는 자료를 참고로 하여 응답을 생성하도록 유도하는 기술이다(Amazon Web Service, 2024). RAG는 LLM의 파인 튜닝보다 훨씬 비용적으로 경제적이다. 이 방법은 파인튜닝을 하는 방법보다 환각 현상 감소에 효과적이고 답변 출력의 일관성을 높이는 데에 더 적합하며(손지원 외, 2024), LLM 자체를 손보지 않으므로 컴퓨팅 자원을 훨씬 적게 요구하는 장점이 있다(Amazon Web Service, 2024).

...

2.2.2 여기에 subsection 제목을 쓰세요

내용을 쓰세요.

수식 작성 예시

$$\text{Cosine 유사도} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

순서 있는 목록 작성 예시

1. 맞춤형 학습: 개인화된 학습 자료 제공
2. 체험 학습: 개인 경험 성찰
3. 잡담(Small-talk): 챗봇에 잡담 기능을 통합하면 사용자를 참여시키고 신뢰를 얻을 수 있는 장점이 있음.

순서 없는 목록 작성 예시

- 교사용 교과서 및 지도서 PDF 파일 내용 데이터셋 28,823개
- 학교 규정 395개

제 3 장

연구 방법

3.1 선행연구에서 수집한 자료의 재범주화 및 비교분석

내용 작성

3.2 챗봇 개발

3.2.1 LLM에 RAG를 적용한 생성형 AI 챗봇 개발 사례

내용 작성

3.2.2 프로토타입 개발 과정

그림 삽입 예시

PDF를 그림으로 넣어도 되고, png 파일을 넣어도 됩니다. 파워포인트로 그린 그림을 PDF로 저장한 그림을 넣는 예시입니다.

표와 그림은 모두 Cref로 언급하면 자동으로 표와 그림을 나누어 번호가 매겨집니다.

챗봇 프로토타입 개발 과정에서 LLM에 RAG를 적용하여 답변을 생성하는 과정을 개략적으로 나타내면 그림. 3.1과 같다.

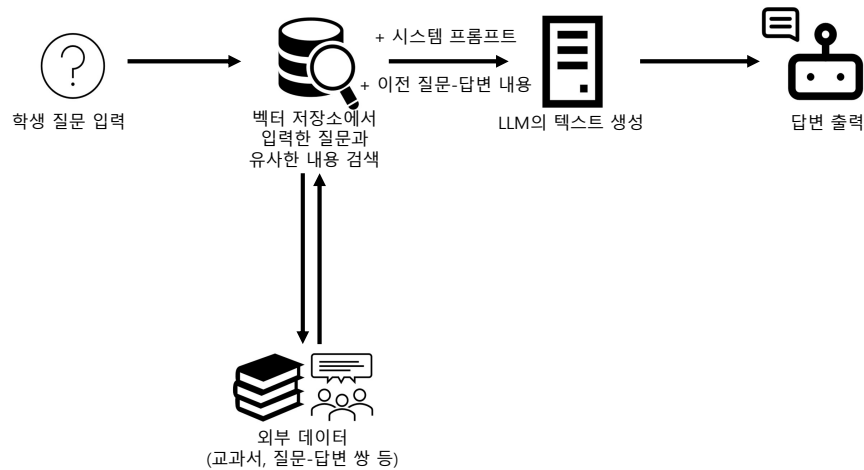


그림 3.1: LLM에 RAG를 적용하여 답변을 생성하는 과정

/figs 폴더에도 그림을 넣어서 사용했습니다. 이 디렉토리는 비워두겠습니다.

subsection도 넣을 수 있어요

...

subsection도 넣을 수 있어요

...

subsection도 넣을 수 있어요

민경모·유준희(2024)가 ... 처럼 쓰면 이/가도 자동으로 바뀝니다.

subsection도 넣을 수 있어요

강조용 글꼴 변화는 `texttt`를 사용하세요.

PDF 파일을 불러온 내용을 내용을 일반적인 문장으로 만들기 위해 `\x`로 시작하는 ... 부분을 제외하고자 진도체크, 보고서 작성하기, 도움 영상, 발표하기, 고르시오, ㄱ, ㄴ, ㄷ, 만들어 보자, 평가하기 등 총 50가지 정도의 문구가 있는 문장을 삭제했다.

...

시스템 프롬프트 구성

컴퓨터 코드 삽입 예시

2024년 9월에 최종적으로 수정한 시스템 프롬프트는 그림. 3.2와 같다.

```
template = '''[BOS]system
당신은 고등학교 과학 선생님입니다. 학생들의 질문에 200 단어 이내의 한국어로만 답변해야 합니다.
중요: 다음 규칙을 반드시 따르세요.
1. 출력 관련 지시사항:
(1) Markdown 을 써서 가독성을 높이세요.
(2) 모든 수식은 무조건 이중 '$' 기호로 감싸서 출력해야만 합니다. 예: $$ 3 \text{ mol } $$
(3) http 나 https 로 시작하는 URL 은 무조건 plain text 로 표현하세요 !
2. 학생이 입력한 질문이 " 안녕하세요 ? ", " 안녕 ? ", " 고마워 ", " 감사합니다 " 와 같은 인사
나 감사 표현이라면:
(1) HISTORY 와 CONTEXT 를 완전히 무시하세요.
(2) 오직 해당 인사에 대해서만 간단히 응답하세요.
(3) 이 경우, ' 생각 '과 ' 답변 ' 구조를 사용하지 마세요.
3. 그밖의 모든 질문에 대해서는:
(1) ' 생각 '과 ' 답변 ' 두 부분으로 나누어 응답하세요.
(2) CONTEXT 와 HISTORY 로부터 답변할 수 있는 경우 가장 우선적으로 참조하여 답변하세요.
(3) 관련 정보가 없는 경우, 자체 지식을 사용하세요.
(4) 답변할 수 없는 경우 " 죄송합니다. 제가 알지 못하는 내용입니다. 라고 " 말하세요.
(5) 고등학교 1 학년 학생이 이해할 수 있는 적절한 용어를 사용하세요.
HISTORY: {history}
CONTEXT: {context}
HUMAN: {question}
[|endofturn|]
[BOS]human
{question}
[|endofturn|]
[BOS]assistant

[|endofturn|]'''
```

그림 3.2: 교사와 학생의 평가를 반영하여 개선한 시스템 프롬프트

표 3.1: 두 가지 학교 맞춤형 과학 질문-답변 챗봇의 사용 기록 및 사용자 평가 내용의 개요

범주	선행연구 1	선행연구 2
자료 수집 기간과 학생의 챗봇 활용 관련 자료 수집 방법	O	O
학생이 주로 챗봇을 활용한 시간과 사용한 기기	O	O
학생의 질문 주제 분류	O	O
학생의 챗봇 사용상 나타나는 특징	O	O
학생의 설문조사 응답 결과	X	O

표 작성 예시

표와 그림은 모두 Cref로 언급하면 자동으로 표와 그림을 나누어 번호가 매겨집니다.

두 가지 학교 맞춤형 과학 질문-답변 챗봇의 사용 기록 및 사용자 평가 내용의 개요는 표. 3.1과 같다.

복잡한 표 작성 예시 1

multicolumn, multirow를 적극적으로 활용하세요. 오류가 가장 적게 납니다.

표 3.2: LLM별 RAG 적용 실험 조건

오픈소스 LLM	실험 조건
LG EXAONE 3.0 7.8B	1. 교사용 교과서와 질문-답변 쌍 모두 활용
	2. 교사용 교과서만 활용
	3. 질문-답변 쌍만 활용
	4. RAG 적용 안 함
Google Gemma 2 9B	1. 교사용 교과서와 질문-답변 쌍 모두 활용
	2. 교사용 교과서만 활용
	3. 질문-답변 쌍만 활용
	4. RAG 적용 안 함

복잡한 표 작성 예시 2

학생 질문을 주제별로 분류한 결과는 표. 3.3과 같다.

표 3.3: 학생 질문을 주제별로 분류한 결과(민경모·유준희, 2022)

대분류	소분류	개수 (비율)
교과 (1,168개, 52.4%)	중3 교과내용 관련	868 (38.9%)
	지필평가	103 (4.6%)
	수행평가	72 (3.2%)
	시험범위	33 (1.5%)
	교육과정	25 (1.1%)
	중1 교과내용 관련	20 (0.9%)
	중2 교과내용 관련	20 (0.9%)
	서술형 시험	16 (0.7%)
	수업	7 (0.3%)
	과제	4 (0.2%)
교과 외 (1,062개, 47.6%)	잡담	503 (22.6%)
	챗봇	187 (8.4%)
	상담	105 (4.7%)
	교사	96 (4.3%)
	타 과목	96 (4.3%)
	인사	42 (1.9%)
	욕설	33 (1.5%)
합계		2,230 (100%)

제 4 장

연구 결과

4.1 section 제목

4.1.1 subsection 제목

subsubsection 제목

subsubsection 제목 subsubsection 아래에 paragraph로 나누어서 작성할 수도 있습니다

...

paragraph 제목 ...

paragraph 제목 ...

제 5 장

결론 및 논의

5.1 결론 및 논의

...

학위논문은 호흡이 길어서 **paragraph***로 나누어서 가독성을 높였습니다 *을 붙이면 차례에 나타나지 않습니다.

paragraph 제목 ...

paragraph 제목 ...

연구의 의의 ...

5.2 연구의 한계

...

5.3 향후 연구 방향

...

마지막으로, 참고문헌은 references.bib에서 본문에 쓰지 않은 것을 굳이 지우지 않아도 됩니다. biber가 알아서 본문에 나온 것만 참고문헌을 작성합니다. 한국어 논문을 참고문헌으로 쓸 경우 references.bib에 keywords=kobib를 넣어주기만 하면 영어 논문과 한글 논문을 구분하여 참고문헌이 자동으로 작성됩니다.

참고문헌

- Amazon Web Service (2024). RAG란? - 검색 증강 생성 AI 설명 - AWS. URL: <https://aws.amazon.com/what-is/retrieval-augmented-generation/>.
- Chang, Yupeng et al. (2024). “A survey on evaluation of large language models”. *ACM Transactions on Intelligent Systems and Technology* 15.3, pp. 1–45.
- Chin, Christine and Jonathan Osborne (2008). “Students’ questions: a potential resource for teaching and learning science”. *Studies in science education* 44.1, pp. 1–39.
- Eshach, Haim, Yair Dor-Ziderman, and Yana Yefroimsky (2014). “Question asking in the science classroom: Teacher attitudes and practices”. *Journal of Science Education and Technology* 23, pp. 67–81.
- 민경모·유준희 (2022). <Doc2Vec 을 이용한 중학교 과학 질문-답변 챗봇 개발 및 학생 질의 분석>. 『The SNU Journal of Education Research』 31.3, pp. 115–145.
- 민경모·유준희 (2024). <인공지능 문장 분류 모델 Sentence-BERT 기반 학교 맞춤형 고등학교 통합과학 질문-답변 챗봇 -개발 및 1년간 사용 분석->. 『한국과학교육학회지』 44.3, pp. 231–248. ISSN: 1226-5187.
- 손지원·김민성·김부건·박상준·표성민·배지훈·이종혁 (2024). <Fine-tuning 과 RAG 를 활용한 특정 도메인 챗봇 제작 방법 비교>. *Proceedings of KIIT Conference*, pp. 341–344.

부록 A

연구에서 사용했던 약어 목록

- API: Application Programming Interface
- ASGI: Asynchronous Server Gateway Interface
- BERT: Bidirectional Encoder Representations from Transformers
- BOS: Beginning of Sentence
- GPT: Generative Pre-trained Transformer
- GPU: Graphics Processing Unit
- IRQA: Information Retrieval-based Question Answering
- LDA: Latent Dirichlet Allocation
- LLM: Large Language Model
- LTS: Long Term Support
- MMR: Maximal Marginal Relevance

- NLP: Natural Language Processing
- NPU: Neural Processing Unit
- RAG: Retrieval-Augmented Generation
- RoBERTa: Robustly optimized BERT pretraining approach
- SLM: Small Language Model
- UUID: Universally Unique Identifier
- VRAM: Video Random Access Memory
- WSGI: Web Server Gateway Interface
- sLLM: Small Large Language Model

부록 B

IRB 결과 통보서와 연구 참여자 모집 문건, 설문지

IRB 결과 통보서와 연구 참여자 모집 문건, 설문지를 첨부하였다.

교사와 학생으로부터 설문지 응답을 받기 위해 연구 참여자 모집 문건과 설문지와 관련하여 IRB의 승인을 받았다.

Abstract

INPUT YOUR TITLE IN ENGLISH

Gildong Hong

Physics Education

The Graduate School

Seoul National University

Recent advances in Large Language Models (LLMs) have enhanced the potential for developing educational chatbots.

...

...

Keywords: Keyword1, Keyword2, Keyword3, Keyword4, Keyword5, Keyword6

Student Number: 2020-30000