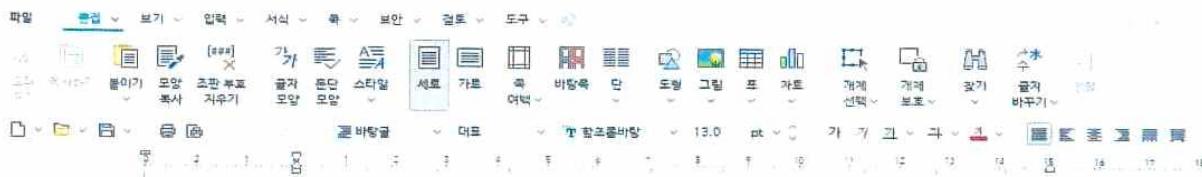
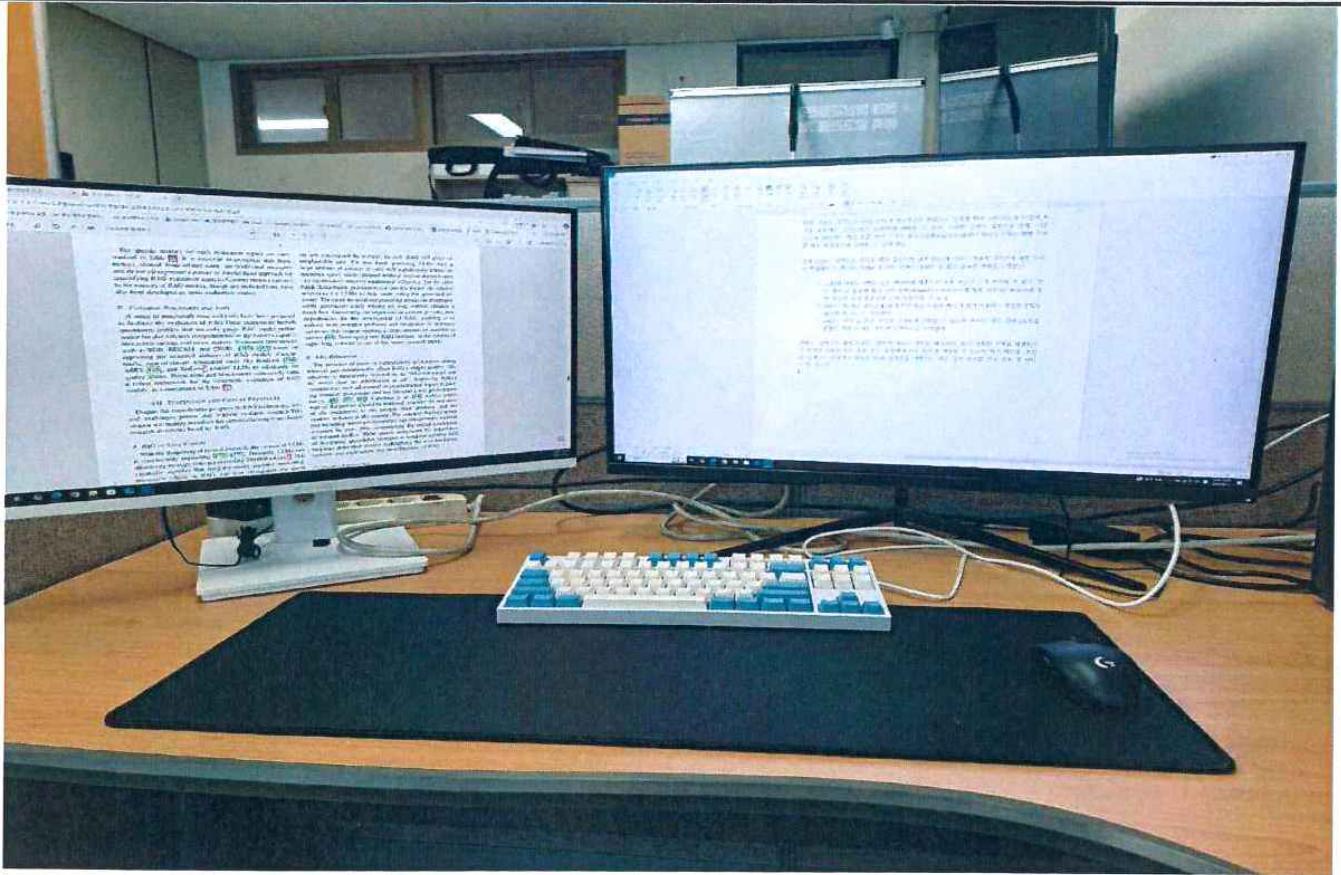


학부생 연구실 인턴 활동보고서

인적사항	소속학과	수학과	성명	김연범
	학번	202000826	학년	4학년
전체 활동 기간	2025. 3. 1. ~ 2025. 8. 31 (6개월)			
해당월 활동 기간	2025. 4. 1. ~ 2025. 4. 30 (1개월)			
활동 내용	<ul style="list-style-type: none"> - 파이썬 기반의 Streamlit 프레임워크를 활용한, RAG 기반 LLM 헌법 Q&A 챗봇 구현 - LLM을 위한 RAG 논문 정독 및 정리 (Retrieval-Augmented Generation for Large Language Models: A Survey) - LoRA(Low-Rank) 논문 정독을 위한 수학적 지식 학습 및 정리 			
활동 성과 및 소감	<p>a) 활동 성과</p> <ul style="list-style-type: none"> - Streamlit 프레임워크의 이해 및 RAG를 활용한 Q&A 챗봇 구현으로 LLM을 위한 RAG의 활용성 및 유용성 이해 - LLM은 어떤 방식으로 동작하고, RAG는 LLM에 어떠한 도움을 주는지에 대한 구체적이고, 전문적인 지식 획득 - Naive, Advanced, Modular RAG 순으로 발전된 RAG 프레임워크의 각 패러다임 분석 - SVD의 개념과 이를 활용한 Rank-r Approximation의 이해 <p>b) 소감</p> <ul style="list-style-type: none"> - Streamlit 프레임워크로 Q&A 챗봇을 직접 구현해 볼 수 있어서 흥미로웠고, 관련 논문을 탐독하며 LLM과 RAG에 대한 전문적인 지식을 얻을 수 있게 돼 뜻깊었다. 			
평가	성실성	업무능력	활용효과	
	(상) 중 . 하	(상) 중 . 하	(상) 중 . 하	
* 활동사진 첨부				
2025년 4월 21일				
보고자 : 김연범 김연범				
담당교수(지도교수) : 김경수 (인)				
충남대학교 소프트웨어중심대학사업단장 귀하				

활동사진(2장)



F. Multi-modal RAG

RAG는 초기의 텍스트 기반 질의응답(QA) 모델에서 벗어나 다양한 모달 데이터를 통합하는 방향으로 발전하고 있다. 이러한 확장은 RAG 개념을 다양한 도메인에서 활용하는 혁신적인 멀티모달 모델을 탄생시켰다.

(1) 이미지 기반 RAG

- a) RA-CM3 : 텍스트 및 이미지의 검색과 생성을 모두 수행하는 최초의 멀티모달 모델
- b) BLIP-2 : 고정된 이미지 인코더와 LLM을 결합하여 효율적인 비주얼-언어 사전 학습을 수행하며, 이를 통해 제로샷 이미지-텍스트 변환을 가능하게 함.
- c) 'Visualize Before You Write' : 이미지 생성을 통해 LM의 텍스트 생성을 유도하는 방법으로, 개방형 텍스트 생성 작업에 유망한 결과를 보임

(2) 오디오 및 비디오 기반 RAG

- a) GSS 방법 : 오디오 클립을 검색하고 연결하여 기계 번역 데이터를 음성 번역 데이터로 변환
- b) UEOP : 외부 오프라인 전략을 활용해 음성-텍스트 변환을 수행하는 end-to-end 자동 음성 인식(ASR) 모델
- c) KNN 기반 Attention Fusion : 오디오 임베딩과 의미적으로 관련된 텍스트 임베딩을 결합하여 ASR 성능을 향상하고, 도메인 적용을 가속화함.
- d) Vid2Seq : 언어 모델에 특화된 시간 마커를 추가하여 사건 경계 및 텍스트 설명을 통합된 출력 시퀀스로 예측하는 모델

(3) 코드 기반 RAG

- a) RBPS : 코드 예제를 검색하여 개발자의 목표에 맞는 사례를 제공하며, 주로 테스트 어설션 생성 및 프로그램 수정을 지원하는 소규모 학습(tail)에서 효과적임
- b) CoK 방법 : 지식 그래프에서 입력 퀴리와 관련된 사실을 먼저 추출한 후, 이를 힌트로 활용하여 지식 그래프 QA 성능을 향상시킴

