

AjouMajorMate

: LangGraph를 활용한 학사정보 안내 챗봇 개발

2025-1 AJOU SOFTCON

이름 신수민

지도교수 강경란

개발동기 및 목적

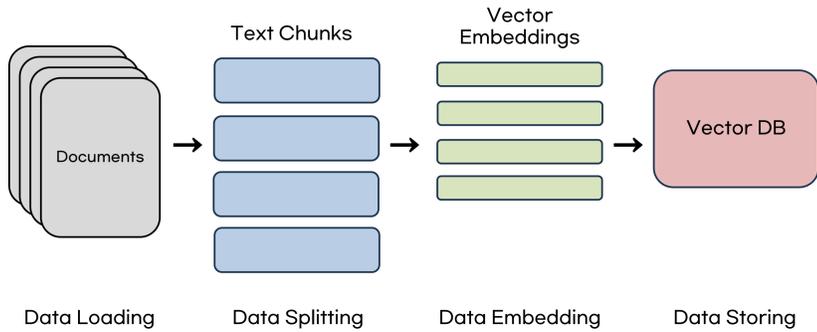
| 대상
자유전공학부 신입생 및 전과·복수전공을 희망하는 아주대학교 학생

| 목적
학과별 요람을 직접 찾아보거나, 학과 사무실에 전화하지 않아도 각 학과별 정보를 손쉽게 얻을 수 있는 대화형 챗봇 서비스를 제공하고자 함.

| 기술적 고려사항
LangChain 기반의 선형 RAG 파이프라인의 한계를 해결하기 위해 LangGraph를 도입하여 그래프 기반으로 각 단계를 유연하게 설계하고자 함.

개발내용

1) Data Indexing



[Data Loading]

- 대상 데이터: 학과별 요람 .pdf 파일 / 학과별 홈페이지 스크래핑 .txt 파일
- 세로 분할 데이터 전처리를 통한 pdf 로드 시 텍스트 의미 보존
- 머리글, 바닥글, 페이지 수 등 검색 과정에서 불필요한 데이터 전처리

[Data Splitting]

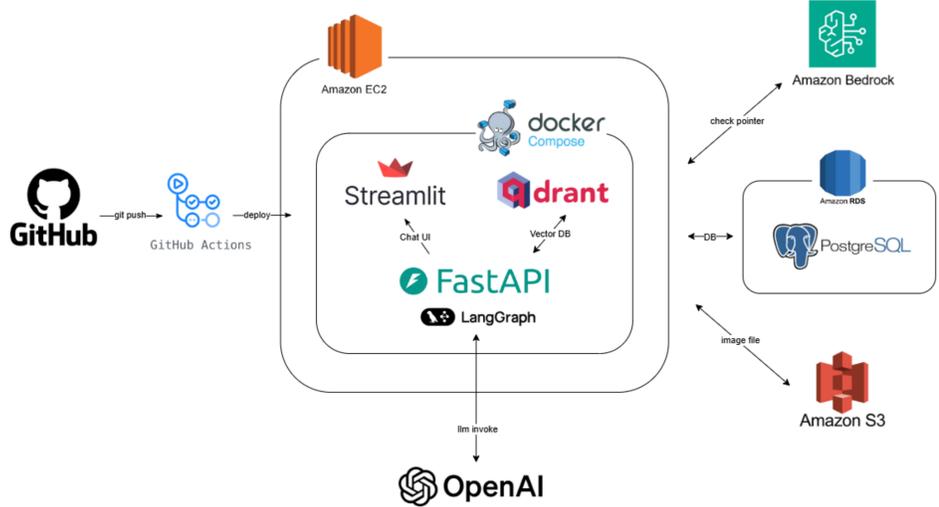
- 요람 내의 수강 과목 데이터 chunking 시 정규 표현식에 기반한 Custom TextSplitter를 사용하여 Chunk 단위별 의미 정확도 향상 및 불필요한 토큰 절감

[Data Embedding & Storing]

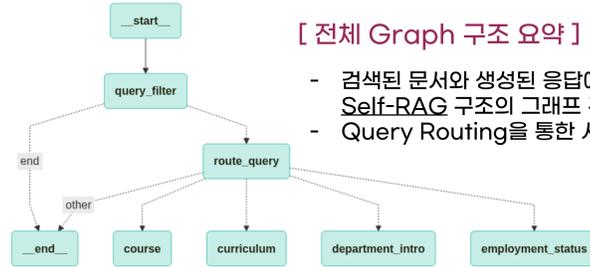
- OpenAI의 text-embedding-3-large 모델을 활용해 chunk 임베딩
- 임베딩 벡터 저장 시 메타데이터를 함께 저장하여 retrieve 정확도 개선 (domain, department, table_id ...)
- Qdrant Vector DB에 Collection을 생성하여 Point 저장

주요기술

[시스템 아키텍처]



2) Data Retrieval & Generation



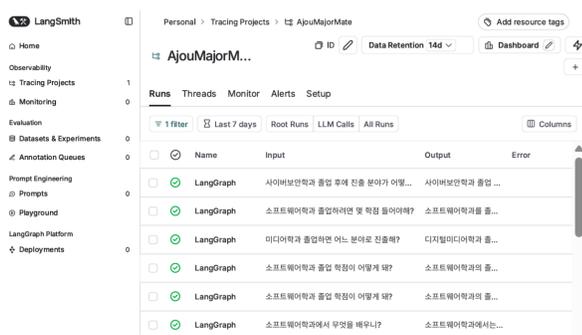
[전체 Graph 구조 요약]

- 검색된 문서와 생성된 응답에 대해 자기 반성 및 자기 평가를 포함하는 Self-RAG 구조의 그래프 구성
- Query Routing을 통한 서버 그래프 분기 처리

[주요 노드 설명]

- query_filter: 질문이 부적절하거나 민감한 내용인지 판단
- route_query: 질문을 분석하여 도메인 분류 후 해당 도메인의 서버그래프로 분기 처리
- extract_department: 질문이 특정 학과에 대한 질문인지 분석한 뒤, 해당 학과명을 메타데이터 필터로 사용
- retrieve: 유사도를 기반으로 문서 검색, 서버 그래프 마다 검색 시 도메인 기반 필터링 적용
- grade_documents: 검색된 문서들이 질문과 관련 있는지 검증
- decide_to_generate: 검색된 문서가 존재하는지 판단 검색된 문서가 없으면 질문 재작성 노드로 이동
- generate: 문서를 기반으로 LLM 응답 생성
- grade_generation_v_documents_and_question: 생성된 응답과 문서를 비교하여 할루시네이션 검증 + 생성된 응답과 질문을 비교해 응답 적절성 검증
- transform_query: 기존 질문이 부정확할 경우, 명확한 질문으로 재작성

결과 및 성과



[문서를 기반으로 답변을 생성하는 챗봇 시스템 구축]

- 도메인 별 임베딩 대상 데이터 관리(저장/삭제) 기능 구현
- Streamlit 기반 챗봇 UI 구현
- LangGraph 기반 Self-RAG 구조의 챗봇 구현
- Query Routing을 통한 서버 그래프 분기 처리로 데이터 Pool을 구분하여 검색 정확도 개선
- LangSmith를 통한 노드 흐름 추적 및 성능 개선
- Docker Compose 및 AWS 기반 서비스 배포

[LangChain 오픈소스 버그 픽스 기여]

- 정규 표현식 기반의 Chunking 구현 중 LangChain TextSplitter 버그 발견 -> 수정 사항 이슈 및 PR 등록
- 공식 PR 반영 되어 컨트리뷰터 등재

오픈소스 URL

<https://github.com/suminnnn/ajou-major-mate.git>

