

응급처치 대응 매뉴얼 조회 및 병원 위치 안내 시스템

작성자: 김지연 작성일: 2024-12-18

구분	ID	서비스(메뉴)	필요 기능	기능 설명	레퍼런스	우선순위
사용자	R-001	응급처치 매뉴얼 조회	응급처치 절차 실시간 제공	사용자의 질의에 맞춰 LLM 모델을 통해 골든 타임 내에 필요한 응급처치 절차를 실시간 제공 긴급 상황에서도 정보 제공의 지연을 최소화하여 즉시 처치 절차를 확인 가능	정보 누리집 응급처치 가이드 pdf, 스포츠 안전재단 응급처치 매뉴얼 pdf, GPT-4omini 모델 문서	1
		병원 위치 조회	위치 기반 병원 검색 병원 위치 시각화	입력한 위치를 기반으로 인근 병원 정보 제공 찾고자 하는 병원의 위치를 지도와 함께 보여줌	공공데이터포털 병원 현황 CSV 파일	1
관리자	A-001	LLM 모델 관리	데이터 전처리 및 관리	응급처치 매뉴얼 PDF 데이터를 텍스트로 변환, 전처리, 임베딩하여 ChromaDB에 저장	GPT-4omini 모델 공식 문서, chromaDB 데이터 저장 및 임베딩 문서	1
			파인튜닝 모델 관리	Fine-tuned GPT-4omini 모델을 기본으로 최신데이터를 학습시키고 모델 업데이트 수행		2
			모델 성능 모니터링	파인튜닝 모델의 응답 정확도 및 성능을 지속적으로 모니터링하고 최적화		
	A-002	데이터 관리	병원 데이터 관리	공공 데이터 CSV파일을 기준으로 병원 정보(위치, 전화 번호 등)를 주기적으로 업데이트	공공데이터포털 병원 데이터, 응급처치 매뉴얼 pdf	1
	A-003	로그 관리	사용자 질의 및 시스템 로그 저장	사용자 Query 기록, LLM 응답 결과 로그를 저장	Django 로그 시스템 문서, Unicorn 로그 관리 문서	1
시스템	S-001	모델 기술 스펙	LLM 모델 파인 튜닝	Fine-Tuned GPT-4omini를 사용하여 응급처치 가이드에 최적화된 모델 제공	GPT-4omini 모델 공식 문서	1
			Retrieval 기반 RAG	Query와 매칭된 문서를 retrieval 후 LLM에 전달해 최적의 응급처치 답변을 생성	ChromaDB 문서	1
	S-002	시스템 최적화	스트리밍 응답	비동기 워커 클래스(gevent)를 활용하여 응답 처리 속도를 최적화	StreamingHttpResponse 기술 문서, Unicorn gevent 비동기 워커 문서	2
			보안 강화	Nginx 설정으로 HTTP 지원, 데이터 암호화 적용 및 서버 접근 제어		
	S-003	서버 관리	unicorn 서버 설정	unicorn을 사용해 django 애플리케이션 실행 및 관리.	Unicorn 공식 문서, Nginx 리버스 프록시 설정 문서, AWS EC2 확장성 설계 문서	1
			Nginx 웹 서버 설정	Nginx를 리버스 프록시로 사용해 unicorn과 연동 proxy_buffering 비활성화로 실시간 스트리밍 지원		2
기타	M-008	시스템 모니터링	상태 확인 및 재시작	시스템 상태 모니터링 및 오류 발생 시 자동 재시작 기능 지원	AWS CloudWatch 공식 문서, Prometheus 시스템 모니터링 문서, Systemd 자동 재시작 설정 문서	2
공통	C-001	디자인 컨셉	직관적 UI/UX	응급상황에 즉시 대응할 수 있도록 간결하고 직관적인 UI/UX 제공	내부 구현	1