

KOS-오더 시스템

자연어 기반 SQL 쿼리 생성 AI 에이전트

Azure OpenAI

Azure AI Search

LangChain

Streamlit

KTDS MS AI 역량강화 과정 MVP 프로젝트
유선오더서비스팀 최지우

프로젝트 개요

문제 정의

- 복잡한 기업 데이터베이스에서 SQL 쿼리 작성의 어려움
- 테이블 구조와 관계 파악에 많은 시간 소요
- 비개발자도 데이터에 접근할 수 있는 도구 필요

대상 사용자

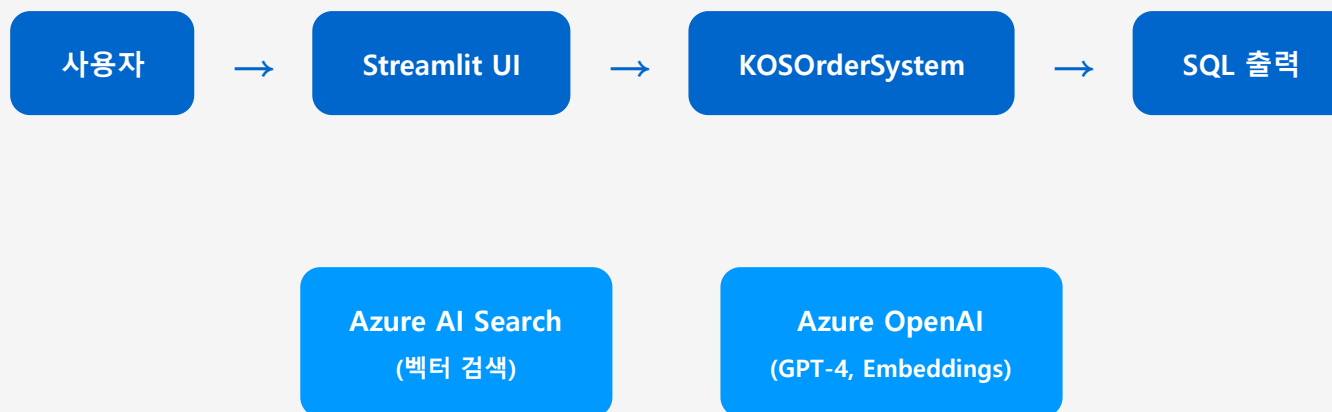
- 데이터 분석가 및 비즈니스 사용자
- SQL에 익숙하지 않은 현업 담당자
- 빠른 데이터 조회가 필요한 의사결정자

솔루션 개요

자연어로 질문하면 AI가 자동으로 SQL 쿼리를 생성

- Azure AI Search를 통한 관련 테이블 자동 검색
- LangChain 기반 구조화된 출력으로 정확한 SQL 생성
- 대화형 인터페이스로 쿼리 수정 및 개선

시스템 아키텍처



주요 컴포넌트



Azure AI Search

벡터 검색으로 의미적으로 관련된 테이블 찾기



Azure OpenAI

GPT-4를 활용한 자연어 이해 및 SQL 생성



LangChain

프롬프트 관리 및 구조화된 출력 처리



Streamlit

직관적인 웹 기반 사용자 인터페이스

처리 프로세스



주요 특징

- 사용자 친화적인 대화형 인터페이스
- AI 기반 자동 테이블 매칭 및 SQL 생성
- 실시간 검증 및 구조화된 출력
- 이전 대화 컨텍스트 유지 및 쿼리 수정 지원

주요 기술 스택 & 특징

기술 스택

- **Python**: 백엔드 개발 언어
- **LangChain**: LLM 오케스트레이션
- **Azure OpenAI**: GPT-4 & Embeddings API
- **Azure AI Search**: 벡터 검색 엔진
- **Streamlit**: 웹 UI 프레임워크
- **Pydantic**: 데이터 검증 & 구조화

주요 특징 및 고려사항

보안

- 환경 변수로 API 키 관리
- Azure 인증 사용

성능

- 벡터 검색으로 빠른 매칭
- 세션 상태 캐싱

확장성

- 모듈화된 구조
- Azure 스케일링 지원

사용성

- 직관적인 UI
- 대화형 인터페이스

핵심 기술 포인트

1. Pydantic을 활용한 구조화된 출력

- LLM 출력을 예측 가능한 구조로 강제
- 타입 검증 및 자동 문서화
- 일관된 응답 형식 보장으로 파싱 오류 최소화

3. LangChain 체인 구성

- 프롬프트 → LLM → 파서 파이프라인
- 재사용 가능한 모듈화 구성
- 에러 처리 및 로깅 통합

2. 벡터 검색 기반 테이블 매칭

- 사용자 질문을 임베딩으로 변환
- Azure AI Search의 k-NN 알고리즘으로 유사 테이블 검색
- 의미적 유사도 기반 정확한 테이블 추천

4. 사용자 경험 최적화

- 대화 컨텍스트 유지로 자연스러운 대화
- 실시간 테이블 컬럼 정보 확인
- 페이징 처리로 많은 테이블도 효율적 탐색

라이브 데모 & 향후 계획

주요 기능 시연

시나리오 1: "어제 발생한 SOIP공사명령송신 내역 시간대별로 조회하고 싶어"

- SOIP 관련 테이블 자동 검색
- 날짜 및 시간 조건 자동 적용
- 시간대별 그룹핑 SQL 생성

시나리오 2: "생성된 쿼리에 할인 ID = 'DC13401' 조건 추가해줘"

- 이전 SQL 컨텍스트 유지
- WHERE 절에 조건 추가
- 수정된 쿼리 즉시 확인

향후 개선 및 확장 계획

성능 최적화

- 쿼리 실행 계획 분석
- 인덱스 추천 기능
- 쿼리 성능 모니터링

기능 확장

- 시각화 자동 생성
- 복잡한 분석 쿼리 지원
- 다중 데이터베이스 연동

엔터프라이즈 기능

- 권한 관리 시스템
- 쿼리 히스토리 및 감사
- 팀 협업 기능

AI 고도화

- 쿼리 최적화 제안
- 자동 데이터 품질 검증
- 비즈니스 인사이트 도출

감사합니다

질문 및 피드백 환영합니다!