**SK네트웍스 Family AI과정 3기  
 모델링 및 평가 시스템 아키텍처**

**□ 개요**

* 산출물 단계 : 모델링 및 평가
* 평가 산출물 : 시스템 아키텍처
* 제출 일자 : 12/30
* 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN03-FINAL-6Team>
* 작성 팀원 : 최연규

|  |  |
| --- | --- |
| 차량추천 | |
| **개요** | 사용자 질문에 기반한 차량 추천 서비스를 제공하기 위해 설계하였습니다. 핵심 목표는 사용자 질문을 이해하고, 데이터베이스에서 적합한 차량 정보를 검색한 후, 최적화된 추천을 제안하는 것입니다. |
| **구성 요소** | * **Input Process**: 사용자 입력을 처리하여 적적한 데이터 형식으로 변환합니다. * **Generate Query**: 전처리된 입력 데이터를 기반으로 데이터베이스에서 검색할 수 있는 쿼리를 생성해 입력한 조건에 맞는 차량 목록을 읽어옵니다. * **Milvus Search**: 사용자 후기가 저장된 Milvus 벡터 데이터베이스에서 사용자의 입력과 가장 유사한 결과를 검색해 차량 목록을 읽어옵니다 * **Rerank Node**: 검색되어진 결과를 재정렬하여 사용자 의도에 가장 적합한 결과를 상위에 배치합니다. * **Suggest Question**: 사용자 질문이 불명확하거나 추가 정보가 필요한 경우, 적합한 후속 질문을 제안합니다. * **Generate Response**: 재정렬된 데이터를 기반으로 사용자 질문에 대한 최종 응답을 생성합니다. |
| **데이**  **터 흐름** |  |
| 차량 매뉴얼 Q&A | |
| **개요** | 차량 매뉴얼 데이터를 기반으로 사용자의 질문에 대한 정확하고 유용한 답변을 제공하기 위해 설계했습니다. |
| **구성 요소** | * **Genesis Check**: 초기 사용자 질문을 분석하여 제네시스 차량 관련 질문인지 판단합니다.. * **Generate History-Based Answer**: 과거 대화 기록을 참조하여 대답이 가능한지 판단하여 답변을 생성합니다. * **Generate Vector Search-Based Answer**: 사용자의 질문을 벡터화하여 Milvus 벡터 데이터 베이스에서 하이브리드 서치 후 리랭크하여 추출된 context를 기반으로 답변을 생성합니다. * **Grade Hallucination**: 생성되어진 답변에 할루시네이션이 있는지 확인합니다. 생성된 답변이 이상이 없다면 점수 측정 노드로 이동하고 이상이 있다면 쿼리 재생성 노드로 이동합니다. * **Calculate Score**: 생성된 답변과 사용자의 질문의 관련성과 정확성을 점수화 하여 평가합니다. 점수가 낮다면 쿼리 재성성 노드로 이동합니다. * **Query Rewrite**: 질문을 답변이 잘 나오도록 재작성하여 개선된 검색 결과를 유도합니다. 재작성된 결과는 다시 **Generate Vector Search-Based Answer**로 전달합니다. * **종료 (End)**: 최종적으로 생성된 답변을 사용자에게 제공하며 프로세스 종료합니다. |
| **데이**  **터 흐름** |  |
| 보험 문의 | |
| **개요** |  |
| **구성 요소** | * **Input Process**: |
| **데이**  **터 흐름** |  |