

SK네트웍스 Family AI과정 7기

모델링 및 평가 LLM 활용 소프트웨어

□ 개요

- 산출물 단계 : 모델링 및 평가
- 평가 산출물 : LLM 활용 소프트웨어
- 제출 일자 : 2025.04.11
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN07-FINAL-5Team>
- 작성 팀원 : 박민혁

개요

- 목적: LLM를 활용한 채용공고 추천 및 모의면접 서비스를 지원하는 소프트웨어 구현

<p>데이터 전처리</p>	<ul style="list-style-type: none"> ● 데이터 수집 <ul style="list-style-type: none"> ○ 면접후기 데이터 <ul style="list-style-type: none"> ■ 내용 : 기업명, 기업코드, 면접 평가, 면접 팁, 면접 질문, 지원 직무, 난이도, 어필역량, 합격여부 등의 정보 ■ 출처 : 캐치 홈페이지 ○ 채용공고 데이터 <ul style="list-style-type: none"> ■ 내용 : 채용공고 제목, 공고 내용 텍스트, 공고 이미지 경로, 채용공고 URL, 기업명, 기업정보, 관련직무, 근무지역, 경력구분, 학력구분 등의 정보 ■ 출처 : 사람인 홈페이지, 사람인 API ● 데이터 전처리 <ul style="list-style-type: none"> ○ 캐치 면접후기 데이터 <ul style="list-style-type: none"> ■ 컬럼 필터링 : 불필요한 컬럼 삭제 ■ 텍스트 정제 : 특수문자 및 기호 (‘, [] 등) 제거, 텍스트 병합 ■ 기업, 직무 테이블 생성 : 면접후기 데이터에서 기업, 직무의 중복값 제거 후 각각의 테이블 생성 ○ 사람인 채용공고 데이터 <ul style="list-style-type: none"> ■ 데이터 중복해소 : 하나의 채용공고에 대한 직무(대분류)가 다수 있어 공고와 직무 분류 ■ 이미지 저장 및 멀티모달 수행 : 공고 본문 이미지를 S3에 적재 후 멀티모달을 수행하여 텍스트 추출 ● 데이터 적재 <ul style="list-style-type: none"> ○ 캐치 면접후기 데이터 <ul style="list-style-type: none"> ■ 대상 : Postgresql ○ 사람인 채용공고 데이터 <ul style="list-style-type: none"> ■ 대상 : Postgresql, Elasticsearch, Chorma ● 사용자 입력 데이터 <ul style="list-style-type: none"> ○ 이력서/자기소개서/포트폴리오 : 사용자가 PDF 업로드 시 텍스트 추출 후 Postgresql 에 적재
<p>기대효과</p>	<ul style="list-style-type: none"> ● 사용자 이력서 및 포트폴리오와 희망 기업, 직무에 따른 개인 맞춤형 질문 제공을 통해 취업 역량 향상 ● 질문의 답변에 대한 논리성, 질문이해도, 직무 전문성, 시간 활용력, 말버릇 습관 등을 분석한 리포트를 통해 사용자의 면접 역량 향상 ● 질문의 답변에 대한 모범답변을 생성하여 면접 역량 향상 ● 사용자 이력서 및 포트폴리오에 따른 채용공고를 추천을 통해 취업 정보 접근성 향상
<p>도구 및 환경</p>	<ul style="list-style-type: none"> ● 클라우드 및 인프라 : GIT, AWS, Docker ● 프론트엔드 : Streamlit ● 백엔드 : Fastapi ● DB : Postgresql, Elasticsearch, Chroma ● LLM 플랫폼 : Openai, Langchain

결론

- 주요 성과
 - 사용자의 이력서 및 포트폴리오와 희망 기업, 직무를 통한 맞춤형 질문 제공
 - 평가 기준별 구체적인 프롬프트 기반 체계적인 모의면접 결과 리포트 제공
 - 하이브리드 검색 기술을 활용한 사용자 맞춤형 채용 공고 추천 시스템 개발
- 향후 계획
 - 데이터
 - Airflow 를 통해 면접 후기, 채용공고 데이터 최신화
 - 기업 경영철학, 기업문화, 지속가능경영보고서와 같은 기업 상세 데이터 확장
 - 직무별 기술 데이터 크롤링
 - 비언어적 표현 분석
 - CNN 모델로 표정분석, 시선처리 등의 비언어적 표현을 분석
 - UI 인터페이스 개선
 - Streamlit 기반 애플리케이션의 소스 코드를 React 기반으로 마이그레이션하여 사용자 친화적 웹 인터페이스 구현
 - AI 에이전트 개발
 - 2D 캐릭터 면접관과 실시간 면접 진행
 - 사용자 피드백
 - 면접 질문, 모범답변, 리포트, 채용공고 서비스에 대한 자세한 피드백을 받아 프로세스 개선