

SK네트웍스 Family AI과정 5기

모델링 및 평가 시스템 아키텍처

□ 개요

- 산출물 단계 : 모델링 및 평가
- 평가 산출물 : 시스템 아키텍처
- 제출 일자 : 2025-01-13
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN05-FINAL-4TEAM.git>
- 작성 팀원 : 박보람, 김지연

개요

AI 기반 인터뷰 시스템은 사용자의 입력 데이터를 처리하고, LLM 기반 AI 모델을 활용하여 인터뷰 질문과 모범 답안 생성, 지원자의 응답 분석, 평가 리포트를 생성하는 구조로 설계되었습니다. 이 시스템은 다음과 같은 주요 목표를 달성합니다:

- 맞춤형 질문 제공.
- 응답의 논리성, 문법, 키워드 적합성을 평가.
- 비언어적 요소 분석 포함.
- 개선 사항을 포함한 상세 리포트 제공.

2.1 데이터 처리 파이프 라인

1. 입력 데이터 처리:

- 사용자가 이력서 폼을 작성하면 S3, PostgreSQL(RDS)에 저장
- 사용자의 음성 답변은 S3에 업로드 후 Whisper 모델을 이용해 텍스트로 변환
- 변환된 음성 답변 텍스트는 PostgreSQL(RDS)에 저장

2. 데이터 저장소:

- AWS S3: 업로드된 원본 데이터(텍스트, 음성 파일, PDF) 저장.
- PostgreSQL: 구조화된 데이터(이력서 텍스트, 음성 텍스트, 질문, 답변, 평가 점수 등) 저장

2.2 AI 모델링

1. 질문 생성 모델:

- PostgreSQL에서 사용자의 이력서 정보 및 기술 스택을 조회
- REST API를 사용하여 파인튜닝 된 LLaMA 모델호출
- 직무 및 기술 스택을 기반으로 맞춤형 질문 생성
- 사용 모델: 파인튜닝 된 LLaMA 모델

2. 응답 분석 모델:

- 사용자가 음성으로 답변 제공 -> S3 저장
- whisper 모델 호출(REST API) -> 음성을 텍스트로 변환
- 변환된 텍스트는 PostgreSQL에 저장
- 응답 분석(GPT-4 기반):
 - i. 논리성, 문법, 키워드 적합성 평가
 - ii. 비언어적 요소(망설임, 말 빠르기, 떨림) 분석
- 사용 모델: Whisper, 파인튜닝 된 LLaMA 모델, GPT-4

2.3 평가 시스템

1. 평가 매트릭스:

- 문법 적합성: 문법적으로 정확한지 평가
- 문제해결능력: 문제를 인지하고 해결하는 과정 설명 여부 평가
- 직무적합도: 기술 질문에 대한 답변의 정확성과 깊이 평가
- 비언어적 요소: 음성의 떨림, 말 속도, 망설임 등을 기반으로 의사 표현 평가

2. 점수 산출 로직:

- AI 모델이 생성한 정량적 점수와 정성적 피드백을 결합하여 최종 점수 계산.

2.4 결과 리포팅

1. 최종 리포트:

- PostgreSQL에서 평가 데이터를 조회
- GPT-4를 사용하여 상세 리포트 작성
- 생성된 리포트는 PDF 형식으로 S3에 저장.
- 워드 클라우드 및 섹션 점수를 시각적으로 포함.

구성 요소	역할
프론트엔드	HTML, CSS, Bootstrap, JavaScript 사용자와 상호작용 (이력서 작성 및 저장, 인터뷰 진행, 결과 리포트 제공)
백엔드	Django, Nginx, Gunicorn <ul style="list-style-type: none">- Rest API를 통한 데이터 요청 처리- Django 기반 데이터베이스 연결
데이터 베이스	PostgreSQL <ul style="list-style-type: none">- 질문, 답변, 점수, 피드백 등 메타 데이터 저장
파일 저장소	AWS S3 <ul style="list-style-type: none">- 원본 데이터(이력서, 음성) 저장- 결과 리포트(pdf) 저장
AI 모델	OpenAI GPT-4, Whisper, LLaMA <ul style="list-style-type: none">- GPT-4: 이력서 키워드 추출, 답변 평가, 리포트 생성- whisper: 음성 데이터 텍스트 변환- LLaMA: 질문 생성, 모범답안 생성
서버리스 컴퓨팅	REST API, Lambda <ul style="list-style-type: none">- REST API를 사용하여 Lambda 실행하고, LLM모델을 호출하는 방식

데이터 흐름

1. 데이터 입력

- **사용자 액션:** 이력서 폼 작성
- **처리 과정:**
 1. 사용자가 작성한 텍스트 데이터는 AWS S3에 저장
 2. 이력서에서 키워드를 추출하여 PostgreSQL(RDS)에 저장

2. 질문 생성

- **사용자 액션:** 인터뷰 시작
- **처리 과정:**
 1. PostgreSQL(RDS)에서 사용자의 이력서 데이터를 조회
 2. 파인 튜닝 된 LLAMA 모델을 호출하여 맞춤형 질문 생성
 3. 생성된 질문은 사용자 인터페이스(UI)를 통해 사용자에게 전달

3. 응답 처리

- **사용자 액션:** 질문에 대해 음성으로 답변
- **처리 과정:**
 1. 사용자가 음성으로 답변을 제공하면 해당 음성 데이터가 AWS S3에 저장
 2. Whisper 모델을 호출하여 음성을 텍스트로 변환
 3. 변환된 텍스트 데이터는 PostgreSQL(RDS)에 저장

4. 응답 분석

- **사용자 액션:** 답변 완료 후 분석 결과 대기
- **처리 과정:**
 1. GPT-4 모델이 미리 정의된 평가지표를 기준으로 답변 자동 평가 진행
 2. Whisper에서 변환한 음성 데이터를 통해 비언어적 요소(망설임, 말 빠르기 등) 분석

5. 평가 저장

- **처리 과정:**
 1. 자동 평가 결과(점수, 피드백)를 PostgreSQL(RDS)에 저장
 2. 저장된 데이터는 이후 리포트 생성에 활용

6. 리포트 생성

- **사용자 액션:** 인터뷰 종료 후 결과 요청

	<ul style="list-style-type: none"> ○ 처리 과정: <ol style="list-style-type: none"> 1. PostgreSQL(RDS)에서 평가 결과를 조회 2. 평가 결과를 바탕으로 강점, 약점, 개선 사항 등을 포함한 최종 리포트 생성 3. 생성된 리포트는 PDF 형식으로 S3에 저장 7. 리포트 제공 <ul style="list-style-type: none"> ○ 사용자 액션: PDF 리포트 다운로드 ○ 처리 과정: <ol style="list-style-type: none"> 1. 사용자가 결과 리포트를 요청하면, 사용자 인터페이스에서 다운로드 요청을 처리 2. AWS S3에서 PDF 파일을 불러와 사용자에게 제공
--	--