

AI 기반 멀티모달 분석을 활용한 모의 면접 분석 시스템 개발 연구

신윤섭, 이우주, 황제윤, 신성진, 김은혜
한국 폴리텍 대학 성남캠퍼스 인공지능소프트웨어 하이테크 과정

개발 배경 및 목적

개발 배경

- 기존 AI 모의 면접 시스템은 면접자의 태도와 응답을 독립적으로 다루어 일회성 피드백에 머무르는 한계가 있다.
- 실제 면접의 인상은 언어/비언어 신호가 결합되어 형성되므로, 단일 신호(텍스트만 또는 표정만) 기반 평가는 맥락을 놓 치기 쉬움.

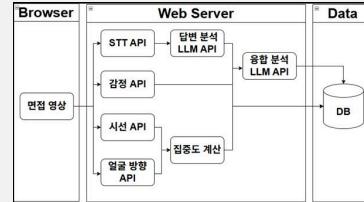
개발 목적

- 멀티모달 시계열 분석으로 면접 중 발생하는 영상·음성 데이터를 동기화하여 집중도·감정·시선·머리움직임 등 핵심 지표를 정량화한다.
- 면접 맥락 기반 피드백을 제공해, 수험자가 자기 인식과 행동 교정을 즉시 수행할 수 있도록 한다.
- 현장 적용성을 고려한 경량 파이프라인을 구축하여, 표준 PC 환경에서도 안정적인 실시간/준实时간 분석을 달성한다.



시스템 사용흐름

- 웹 기반 플랫폼으로 브라우저를 통해 접속하여 로그인 후 면접 유형(실전 면접, 모의 면접)을 선택
- 자기소개서를 업로드하여 개인화된 면접 문항 생성하고 연습할 면접 문항을 선택
- 카메라 및 마이크 테스트 이후 캘리브레이션을 통해 분석 기준점을 설정
- 각 문항에 대해 1분 이내의 모의 면접을 진행
- 면접 회차 전체에 대한 분석 및 문항별 감정, 집중도, 답변 분석을 확인하여 피드백
- 반복 연습을 통해 이전 연습 결과와 현재 연습 결과의 비교 분석을 확인하고 면접 실력 향상을 유도

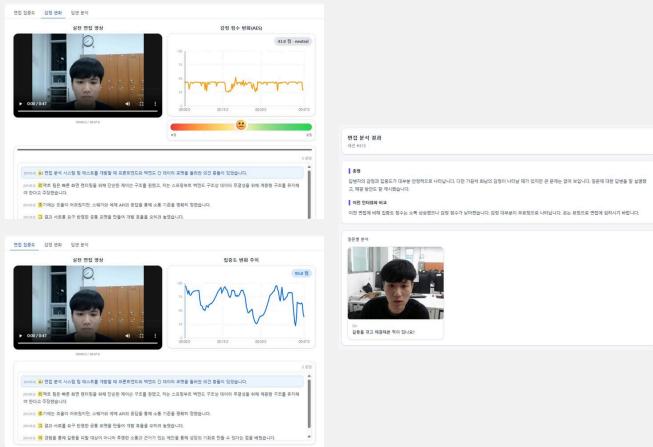


시스템 아키텍처

주요 기능

감정 융합 분석

- STT(Speech-to-Text)를 통해 답변 시계열 데이터를 추출하고, 감정 추론 모델로 도출한 단위 프레임 단위의 감정 데이터와 연계하여,
- 최종적으로 면접 영상과 답변내용, 감정 지수 그래프를, 시간을 매개로 동기화하여 보여줌으로써 어느 시점에 어떤 답변을 하는지, 어떤 감정으로 비춰졌는지 파악할 수 있음



집중도 융합 분석

- 마찬가지로 STT로 추출된 답변 시계열 데이터를 활용, 시선과 머리 방향 데이터를 통해 산정한 집중도 지수를 기반으로
- 면접 영상과 답변내용, 집중도 지수 그래프를 시간 매개로 동기화하여 보여줌으로써 특정 시점의 답변내용과 집중도 지수를 확인할 수 있음

면접 연습 결과 비교 분석

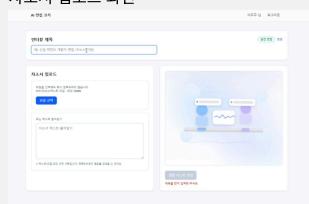
- 직전 면접 분석 결과와 현재 면접 분석 결과를 LLM(DeepSeek R1) 모델에 입력, 비교하여 집중도, 감정, 답변 분석 결과에 대한 개선 정도를 평가하여 표시함

결과

메인화면



자소서 업로드 화면



집중도 분석 화면



답변 분석 화면



질문 선택 화면



감정 분석 화면

