

SK네트웍스 Family AI과정 5기

데이터 수집 보고서

□ 개요

- 산출물 단계 : 데이터 수집 및 저장
- 평가 산출물 : 데이터 수집 보고서
- 제출 일자 : 2025-02-12
- 깃허브 경로 : <https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN05-FINAL-4TEAM.git>
- 작성 팀원 : 박보람

데이터 수집 목적	본 데이터 수집의 목적은 대규모 언어 모델(LLM)의 질문과 답변 생성 성능을 높이기 위함이다. 이로써 실제 면접과 유사한 환경을 조성하고, 고품질의 모의 면접 서비스를 제공하고자 한다.
데이터 수집 방법	<p>3.1 사람인 데이터 (웹 크롤링)</p> <ul style="list-style-type: none">- 사람인 웹 사이트에서 제공하는 백엔드, 프론트엔드, ai 엔지니어 신입 공고에서 직무명, 회사명, 담당 업무, 자격 요건 등을 크롤링한다- 백엔드 신입, 프론트엔드 신입, AI 엔지니어 신입 <p>3.2 평가 지표 데이터 (PDF)</p> <ul style="list-style-type: none">- 인공지능을 활용한 비동시적 면접 연구의 현황과 과제- 음성 신호를 이용한 감정 인식- LLM 기반 CHAT GPT를 활용한 컴퓨터 분야 면접 준비용 (AI 모의 면접 시스템의 설계 및 구현에 대한 연구)
수집 데이터(요약)	<p>3.1 사람인 데이터</p> <ul style="list-style-type: none">- 공고별 직무명, 회사명, 담당 업무, 자격 요건에 대한 내용을 담고 있다 <p>3.2 평가 지표 데이터</p> <ul style="list-style-type: none">- 각종 평가지표 선정 배경에 대한 내용을 담고 있다