

머신러닝을 통한 자기소개서 분석: 길JOB이

고현경 김희주 김혜인 이선희 이수민

길JOB이 프로젝트

참고 영상



취업도 '고액'코디

한 취업학원에서는 3개월
268만원의 비용을 제시했다.



단순히 '믿어라'

전문적인 분석이 아닌 단순히
'믿어라'라는 방식의 분석이다.

자기소개서 작성의 어려움



200만원

자기소개서 작성 관련
취업컨설팅 교육 비용

자기소개서 첨삭, 컨설팅에 들어가는
비용이 과도하게 높은 추세



85.3%

취준생 85.3%
'자소서 작성 어렵다'

자기소개서 작성 경험이 있는 경력직
직장인 85.3%가 자기소개서 작성에
어려움을 겪었다고 대답.

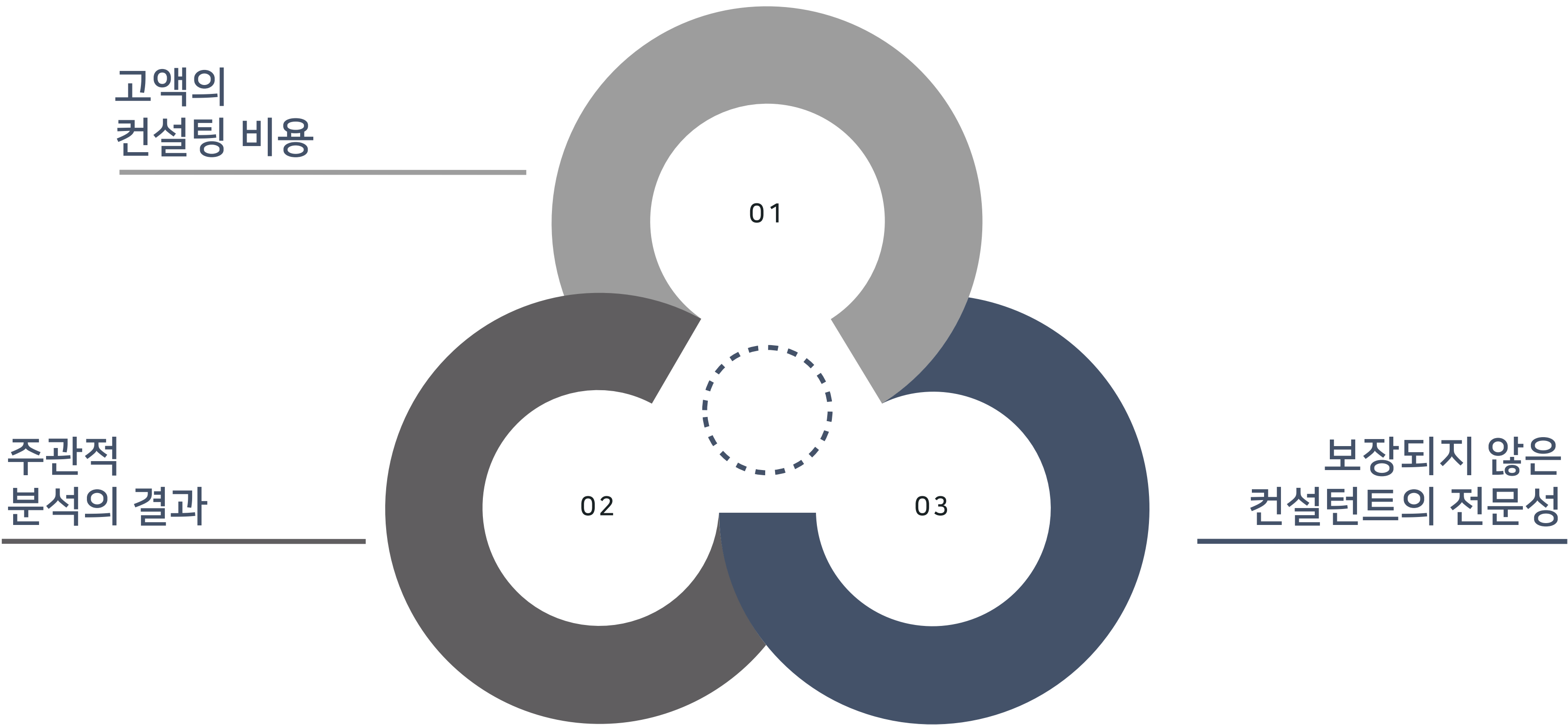


242분

취준생 평균 자기소개서
작성 소요시간

많은 취준생들이 자기소개서를 한번
작성하는데 상당한 시간을 들이는 것으로
나타났다.

현재 취업컨설팅 시장의 한계



길JOB이 프로젝트

길JOB이의 해결책

저렴한
컨설팅 비용

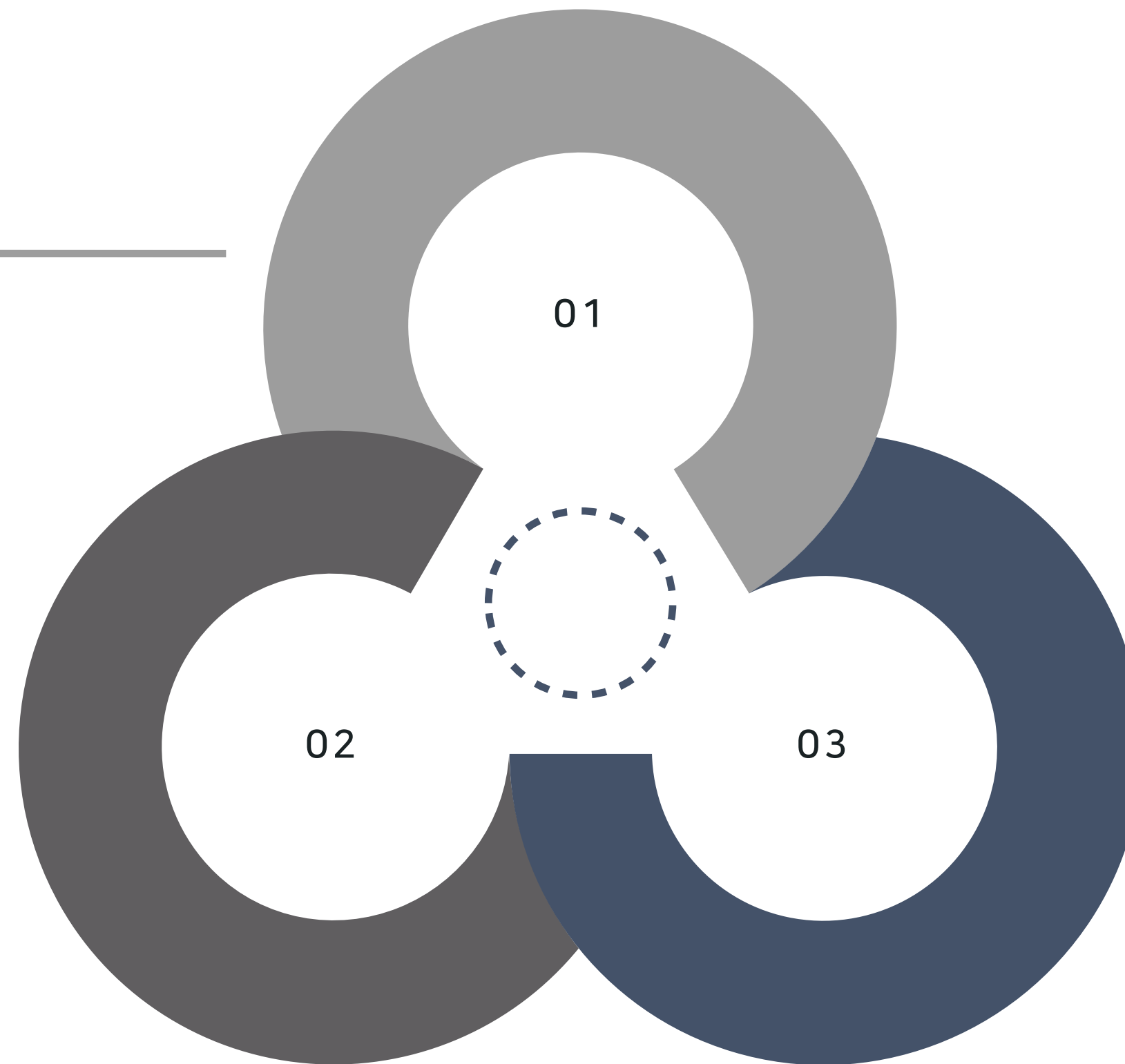
01

객관적
분석의 결과

02

빅데이터를 기반으로한
신뢰도 있는 결과

03



길JOB이 프로젝트

길JOB이의 목표



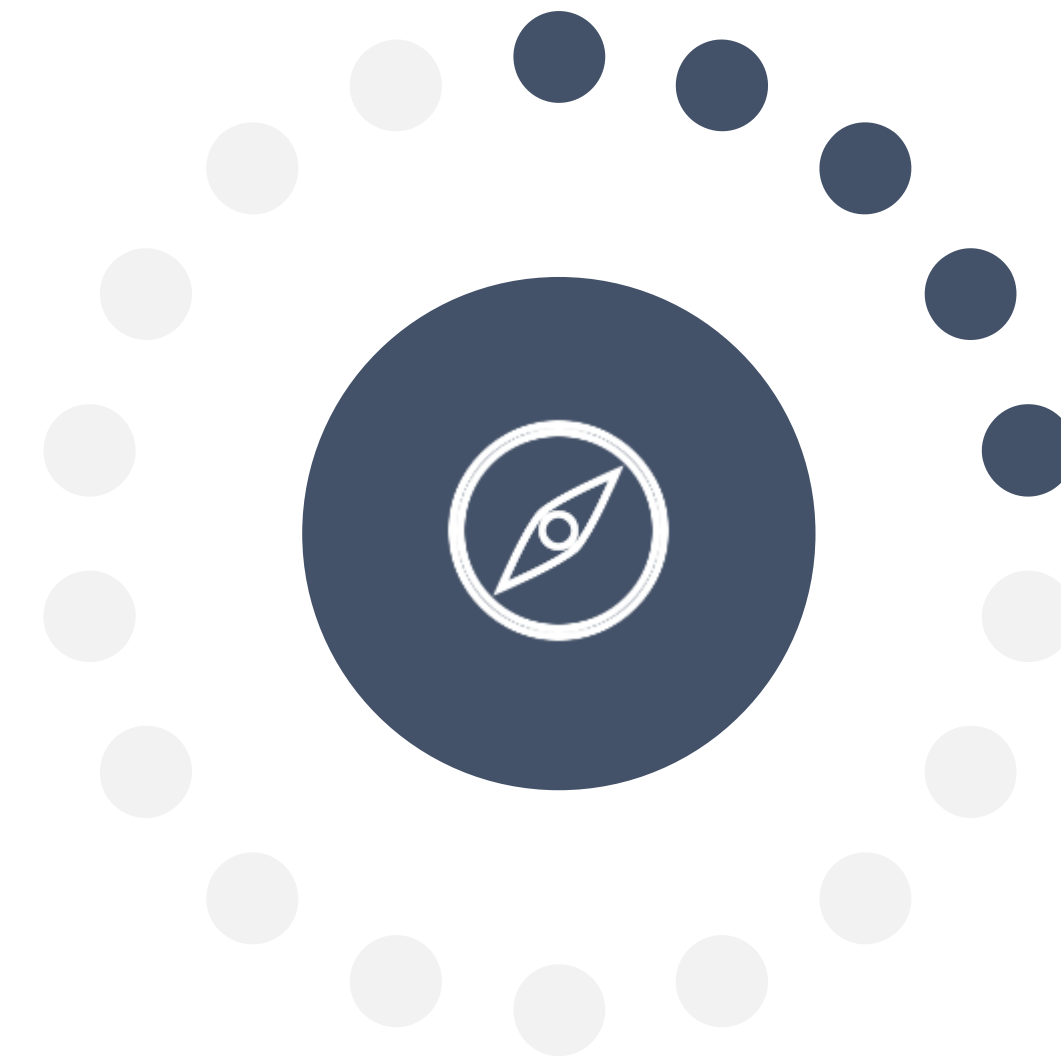
객관적인 평가

자기소개서에 대한 객관적인
평가가 가능하다.



기업 이해도 상승

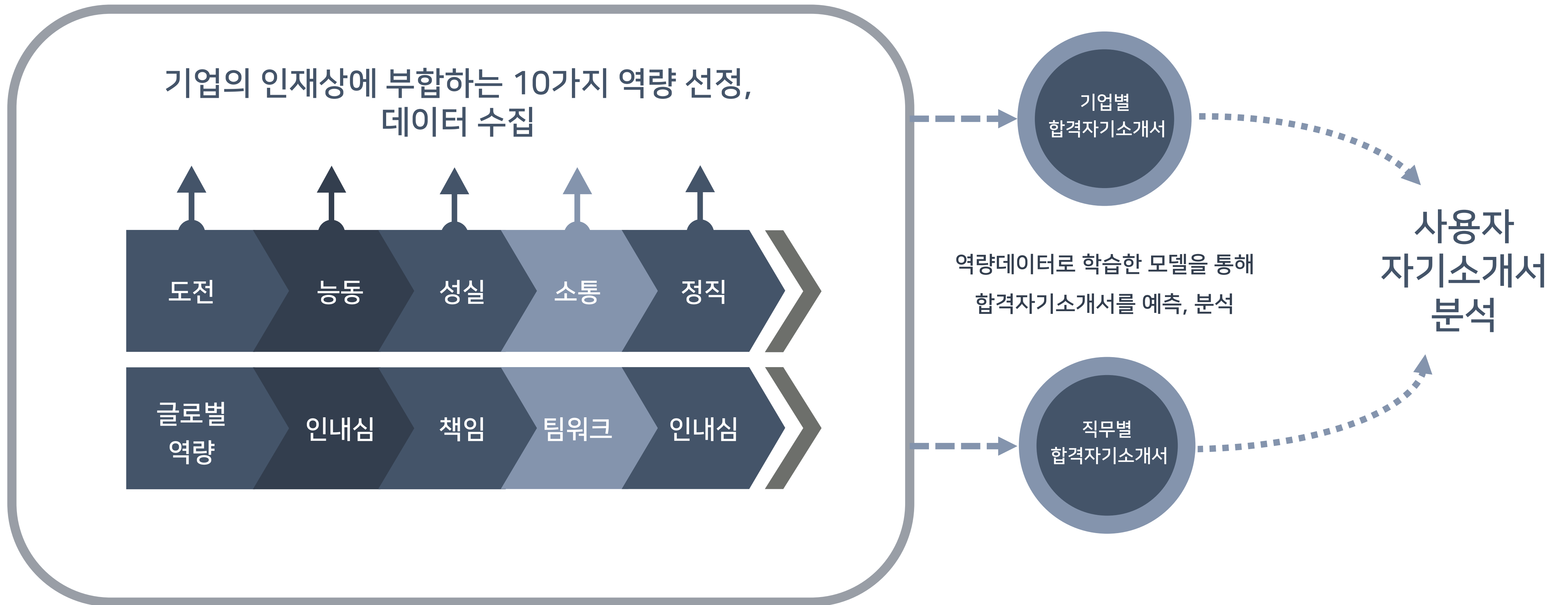
취업을 목표로한 기업에 대한
이해도가 상승한다.



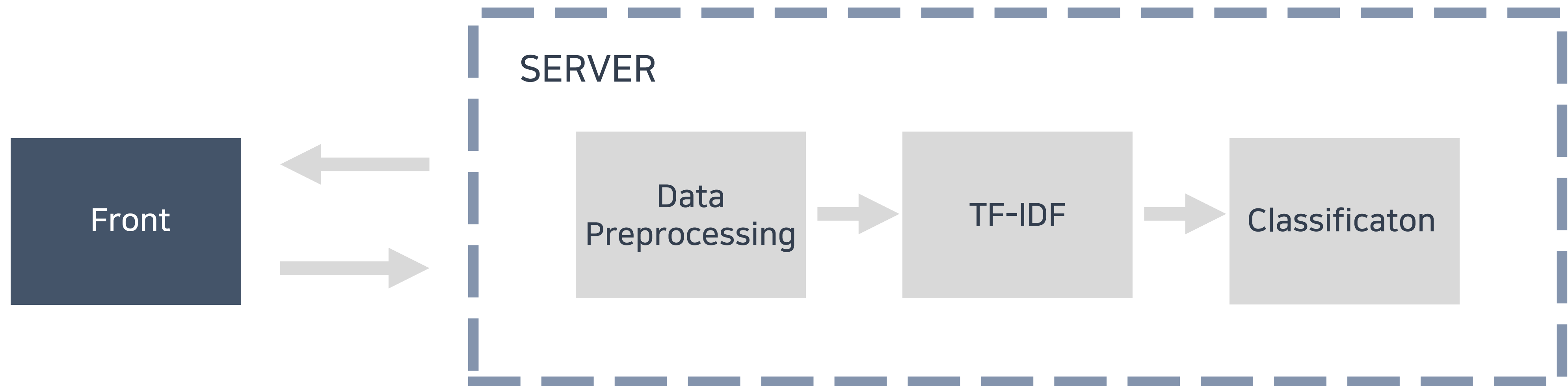
가이드라인 제시

자기소개서 작성에 대한
가이드라인을 제시한다.

| 길JOB이의 기본적인 구성



| 개발과정



Front

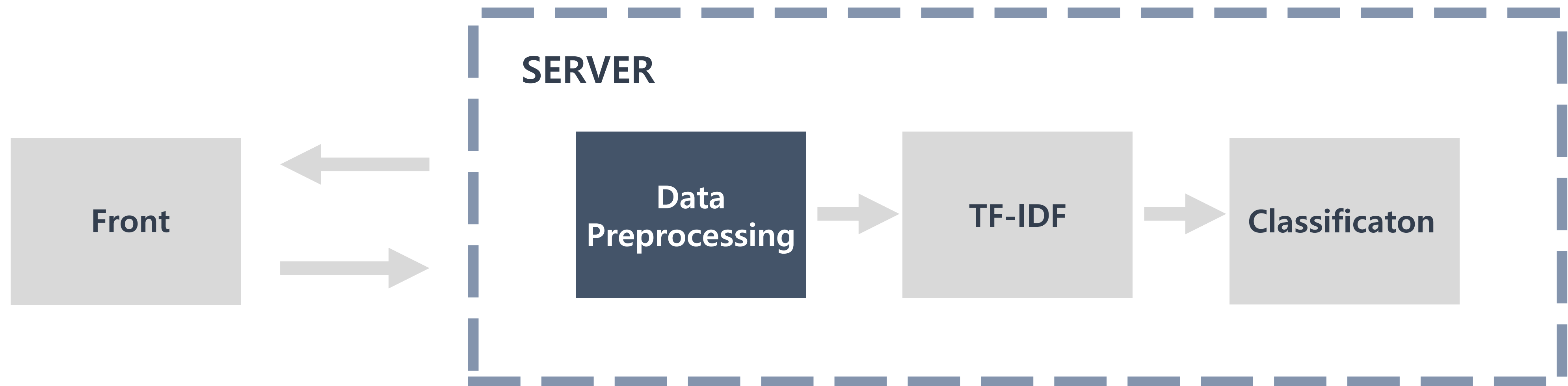
사용자의 자기소개서를 **웹 서버**로 전송.

자기소개서를 **자연어처리 서버**로 전송.

전송 및 결과 수신은 Synchronous 통신으로 구현.

분석 결과를 수신받아 결과페이지를 생성.

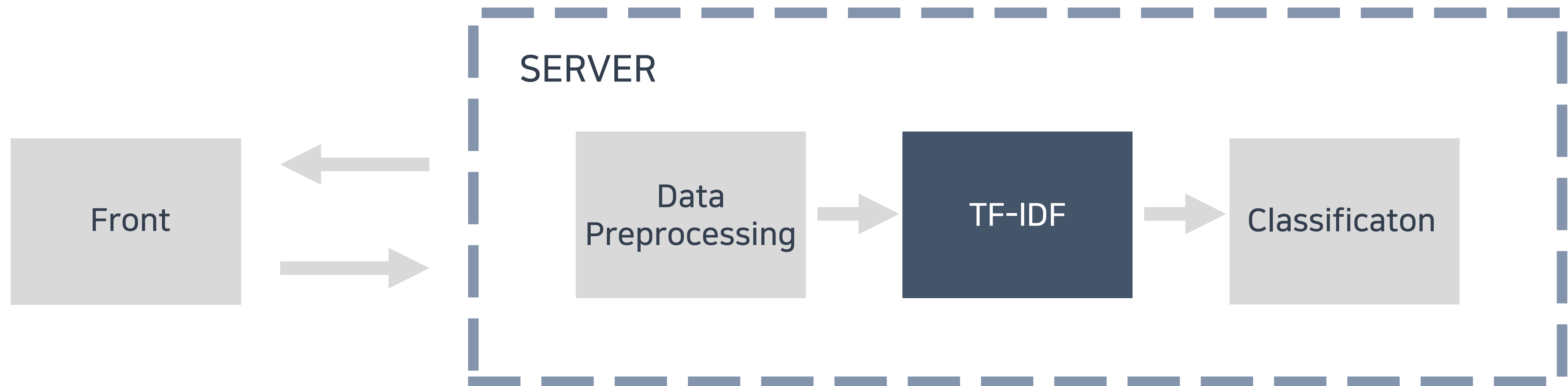
| 개발과정



Data Preprocessing

수집된 데이터를 분석할 수 있도록 정제하는 단계.
한글 전처리 라이브러리 KoInpy 활용.
데이터를 품사별로 분리한 후, 필요한 품사만을 추출.
추출한 단어들을 정규화, 어근화, 불용어제거 순으로 정제.

| 개발과정



TF-IDF

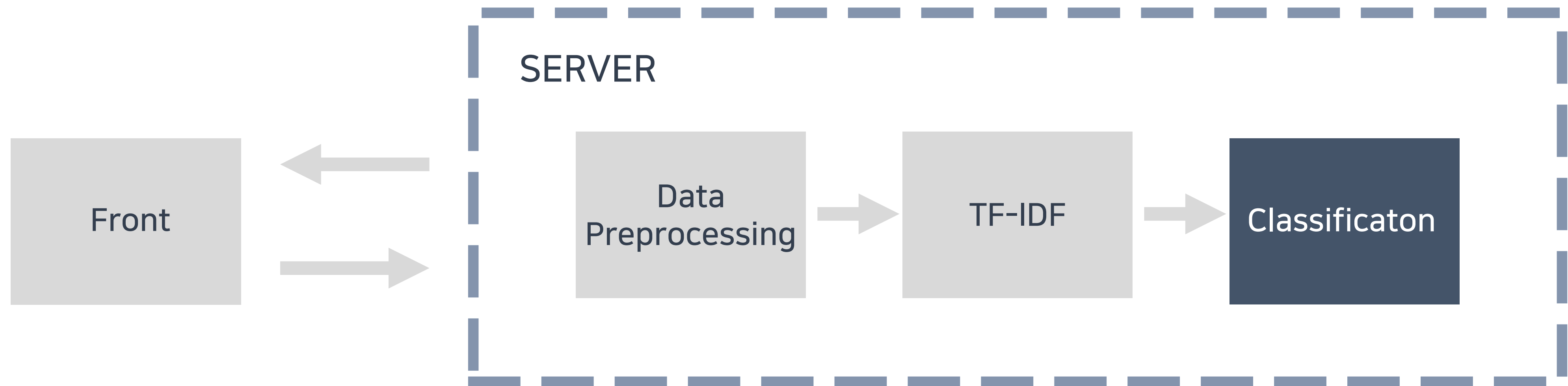
단어의 중요도를 나타내기 위해 TF-IDF방식을 사용.

어떤 단어가 특정 문서 내에서 얼마나 중요한 것인지를 나타냄.

한 문서 내에에서의 단어 출현 빈도와 문서집합 전체에서의 단어 출현빈도 역수를 가지고 단어의 중요도를 나타냄.

TF-IDF값을 사용해 텍스트 데이터를 수치로 변환.

| 개발과정

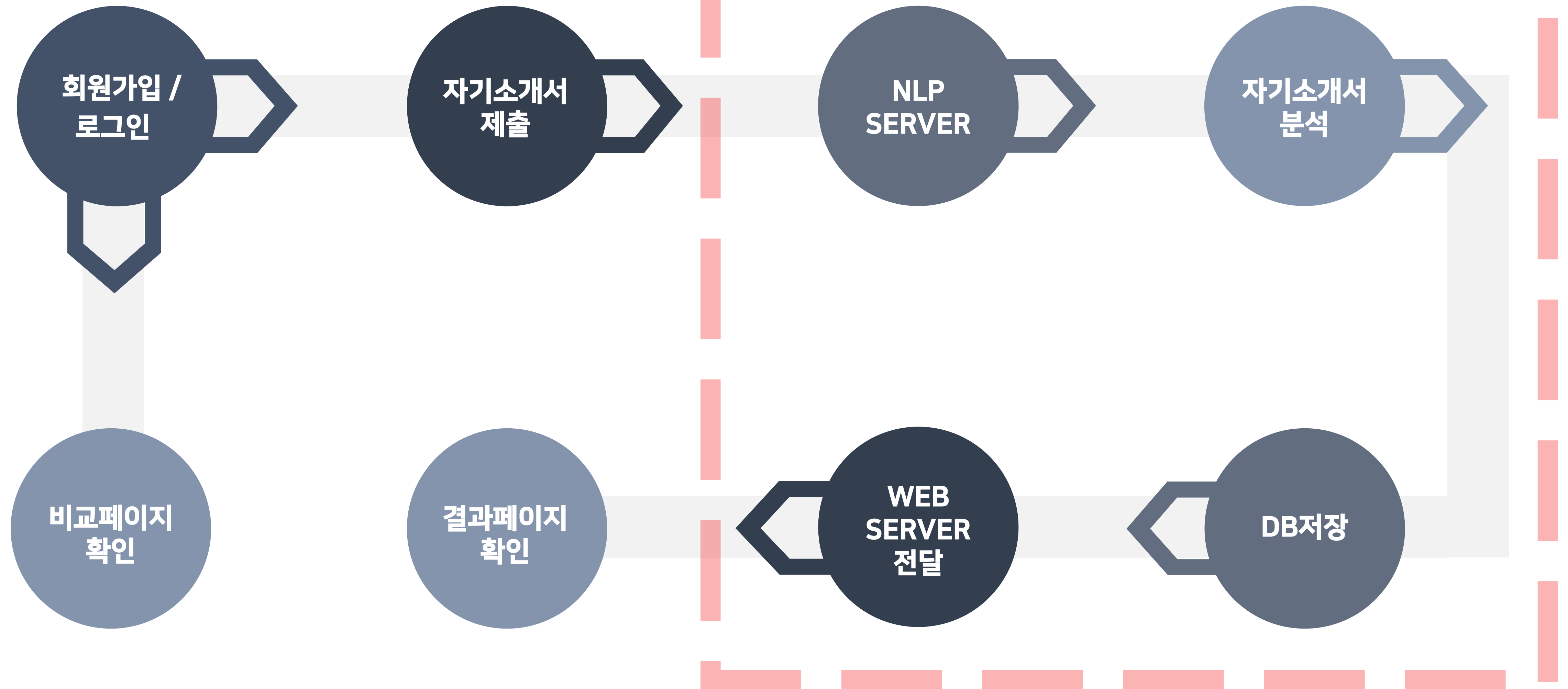


Classification

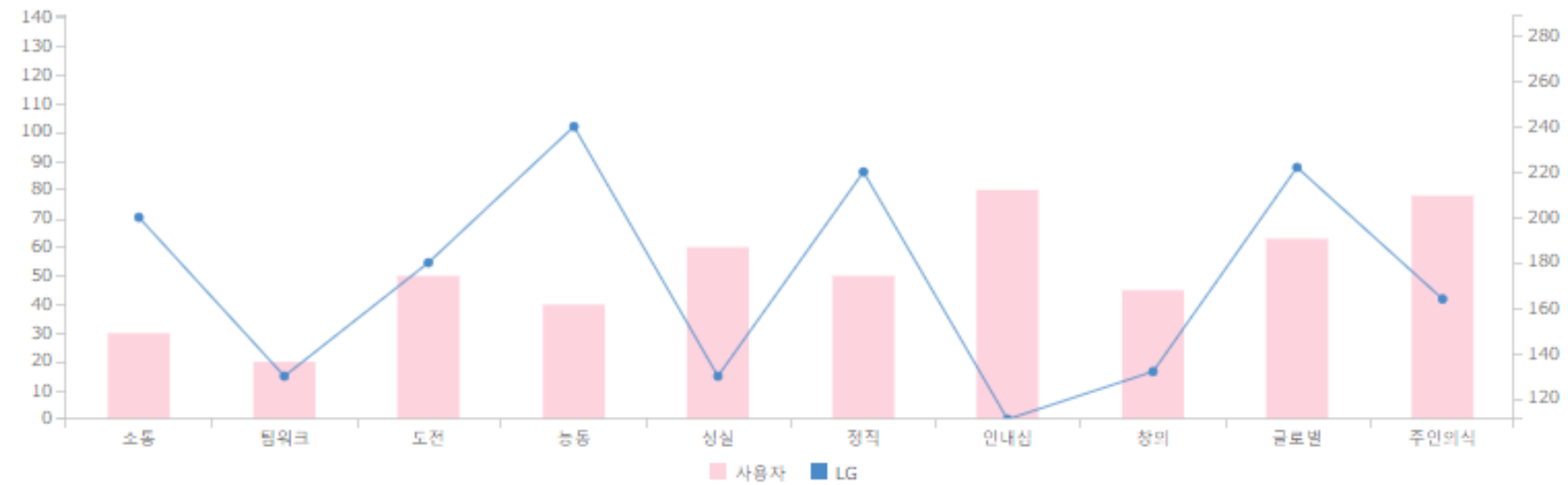
벡터화된 데이터를 사용해 지도학습으로 모델 생성.
다중클래스로 분류된 데이터에 대해서 입력데이터를 각 클래스로 분류.
여러 모델 테스트 후 서포트 벡터 머신(SVM)으로 결정.

| 시스템 진행순서

REST API



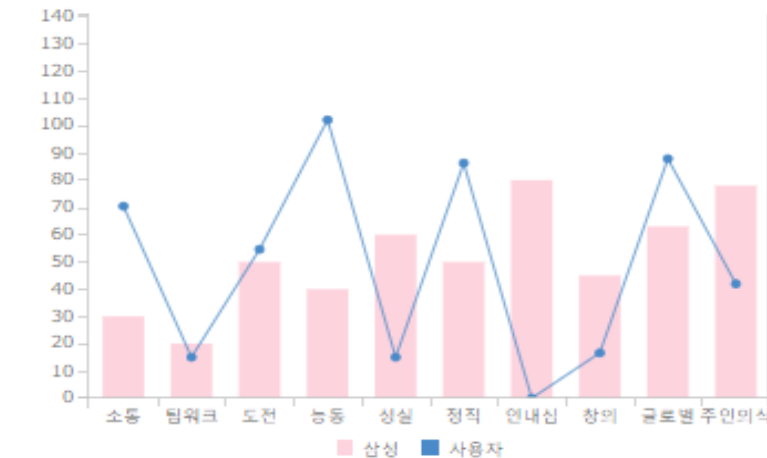
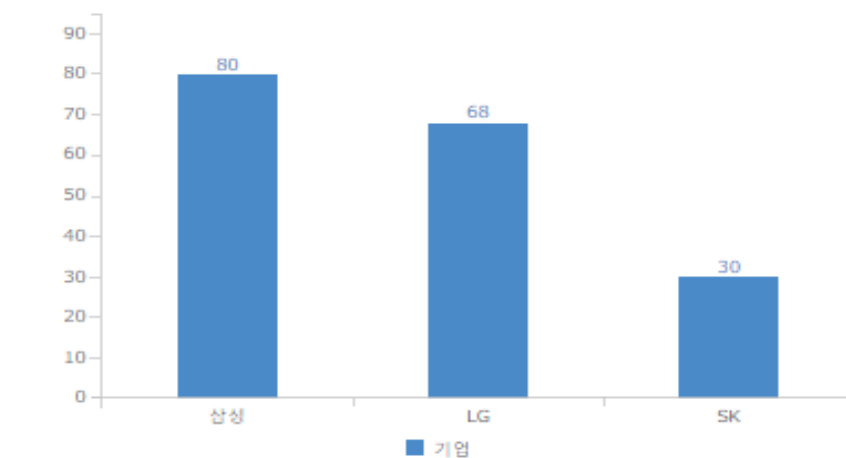
| 결과물 - 결과 페이지 그래프



[선택 기업 핵심역량 그래프]

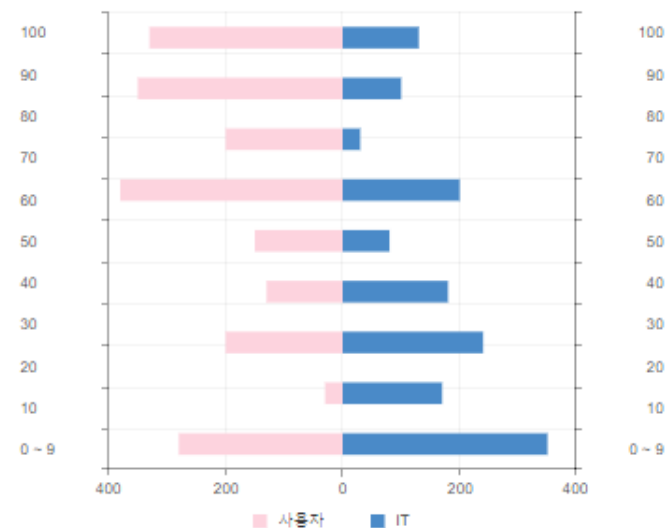
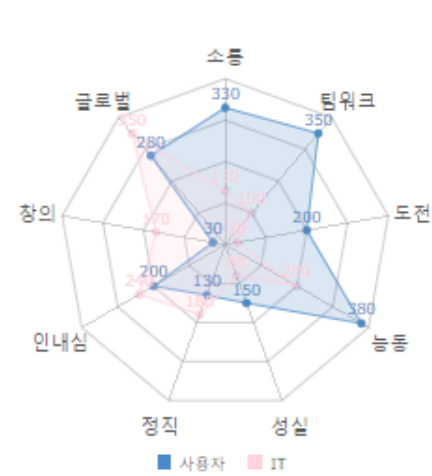
선택한 기업과 사용자의 자기소개서가 핵심 역량 별로 얼마나 일치하는지 보여주는 그래프이다.

[기업별 적합도 그래프]



[추천 기업 순위 그래프] [1순위 기업의 핵심역량 그래프]

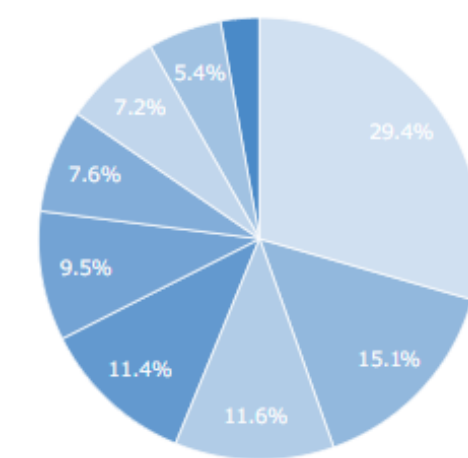
사용자의 자기소개서와 유사한 기업을 3순위로 보여준다. 오른쪽 그래프는 해당 기업과 사용자의 역량 데이터를 보여주는 그래프이다.



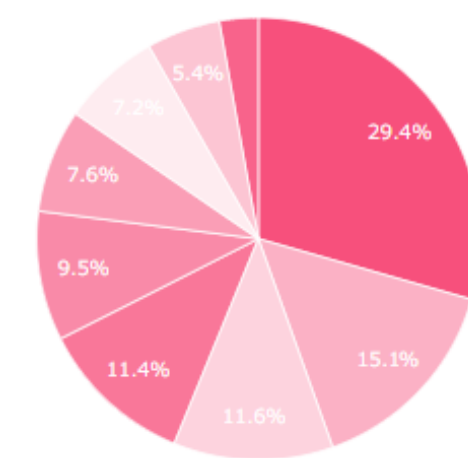
[선택 직무 핵심역량 그래프]

선택한 직무와 사용자의 자기소개서가 핵심 역량 별로 얼마나 일치하는지 보여준다. 오른쪽 그래프는 선택한 직무에서 상위 5개의 핵심역량을 뽑아 왼쪽 그래프의 수치를 한눈에 알아볼 수 있도록 나타낸다.

Q1. 당신의 성장 배경은?



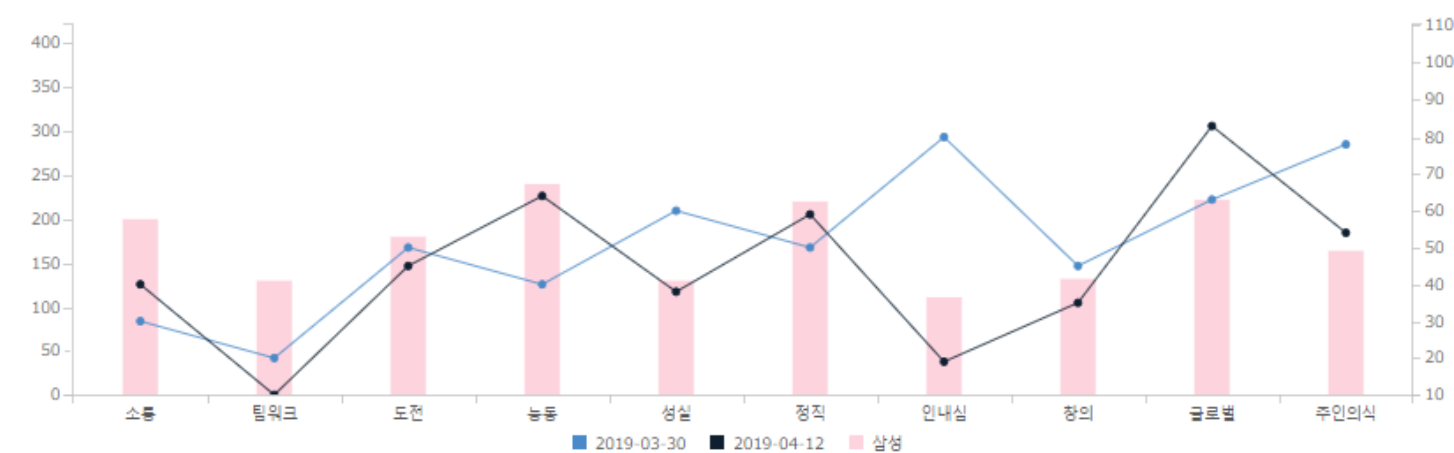
Q2. 당신이 삼성에 지원하게 된 계기는?



[문항별 핵심역량 그래프]

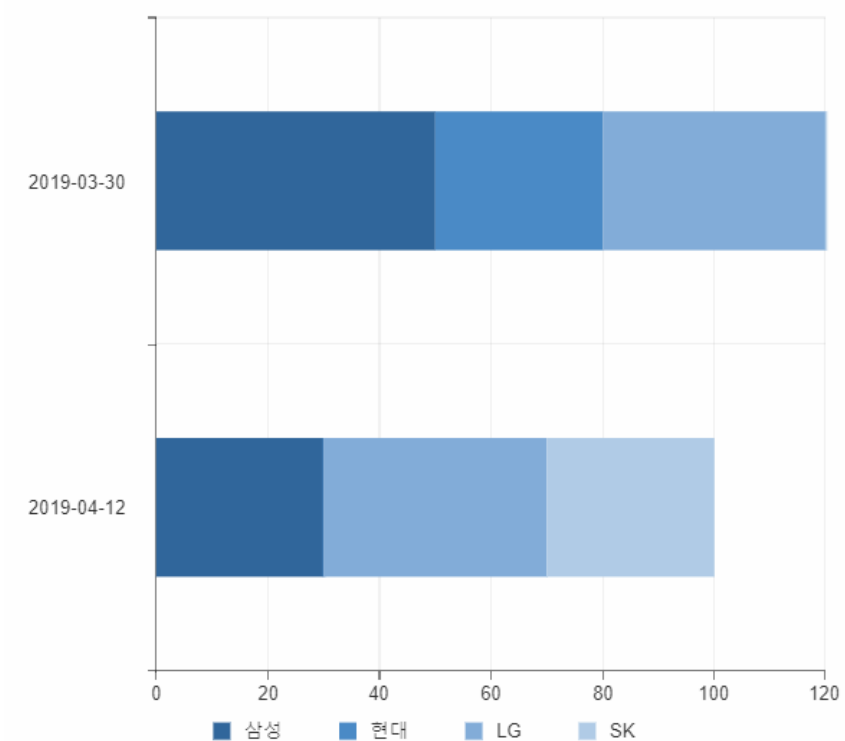
문항별로 직무 별 핵심 역량을 분석한 결과이다.

| 결과물 - 비교 페이지 그래프



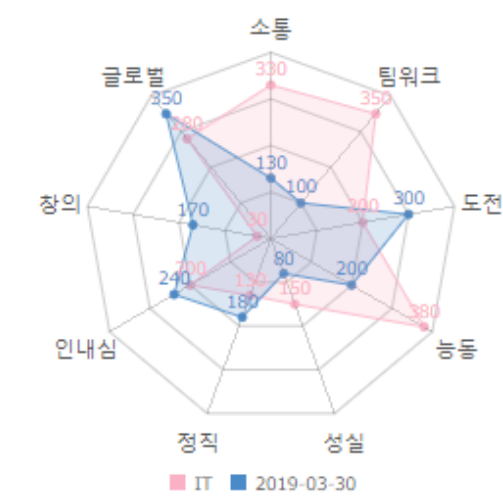
[선택기업 핵심역량 그래프]

선택 기업에 대한 핵심역량 분석 결과를 사용자가 선택한 자기소개서 끼리 비교하여 보여주는 그래프이다



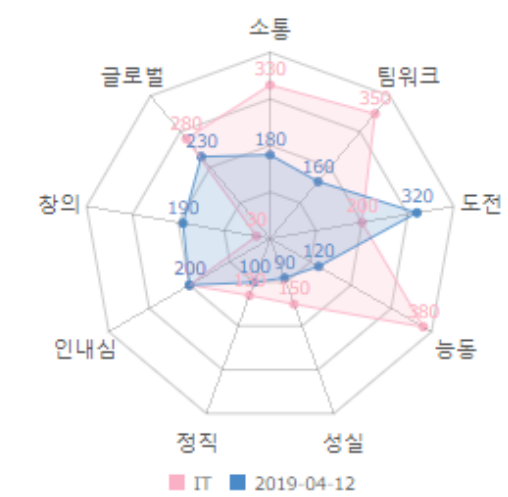
[기업별 적합도 그래프]

사용자의 자기 소개서가 어느 기업과 적합한지 비교 할 수 있도록 자기소개서 별로 나 타낸다.

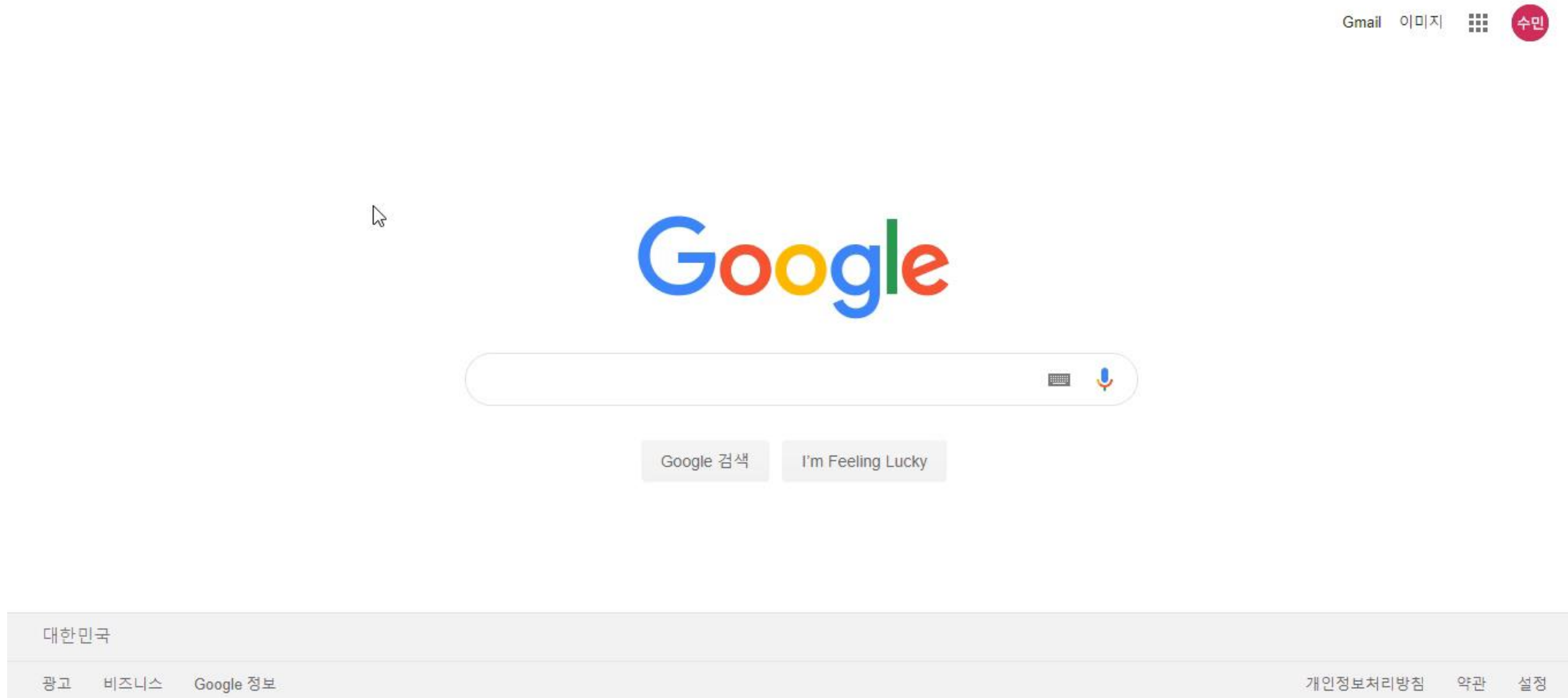


[직무별 핵심역량 그래프]

직무에 대한 핵심 역량 분석 결과 그래프를 날짜별로 각각 보여주어 사용자가 결과를 비교할 수 있도록 해준다.



| 결과물 - 시연동영상



| 기대효과 및 활용방안



기대효과

길JOB이는 자기소개서를 분석하여 사용자의 강점이나 약점을 객관적으로 파악 할 수 있습니다. 기존의 탐색 시스템과는 다르게 객관적인 분석이 가능하며 비용이 저렴하고 시간이 적게 소요됩니다. 사용자가 원하는 직무 뿐 아니라 기업의 인재상, 합격 자기소개서 등으로 분석된 데이터를 바탕으로 사용자의 자기소개서가 어느 기업에 적합한지 알려주어 취업 준비 시 참고할 수 있습니다.



활용방안

추후 직접 기업입장에서 직접 이 기술을 이용해 지원자들의 역량을 평가하는 데 활용할 수 있도록 하는 시스템으로 발전할 수 있으며, 더 나아가 길JOB이가 리쿠르팅 시스템으로 활용되어 기업과 지원자의 교두보 역할을 할 수 있습니다.



THANK YOU

고현경 김희주 김혜인 이선희 이수민