제 2회 고용노동부 공공데이터 활용 공모전

일 찾으러 다니는 건 그만! 고용아 잡(Job)줄래?

불필요한 검색을 줄여주는 채용공고 기반 추천서비스





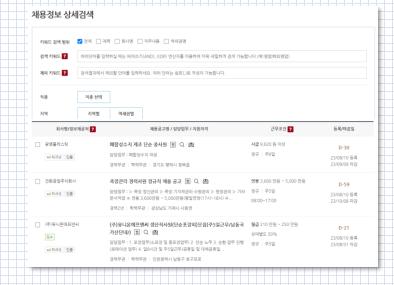
CONTENTS



[90s~00s]



[00s~10s]



[20s~]





- AI일자리추천은 회원님에게 적합한 일자리 및 훈련 등의 정보를 매일 선별하여 제공하고 있습니다.
- 현재 AI일자리추천 채용정보는 기존 알고리즘과 새로운 AI 알고리즘을 운영중에 있습니다.

화원님의 이력서 및 구칙활동 정보와 직업훈련 수강 이력 등을 추천 기준으로 활용합니다. 추천기준 자세히 보기 회원님과 유사한 특징들을 분석하여 회원님에게 더 풍부하고 다양한 고용정보를 제공하고 있습니다.

채용	정책	기업	훈련	자격
6	0	5	8	5

회원님의 이력서를 직무역량을 중심으로 구체적으로 작성하시면 입자리 매칭 정확도를 높일 수 있습니다. - 기업의 구인공고나 구칙자의 이력서가 변경된 경우에는 익일 반영되니 이 점 유의하시길 바랍니다. 추천 정보는 매일 업데이트 되므로 관심있는 정보는 스크랩하시기 바랍니다. 이력서 관리

심리검사

찾는다

찾는다

받는다

[찾으러 다니는 시대에서 가져다주는 시대로]



[Product Vision]

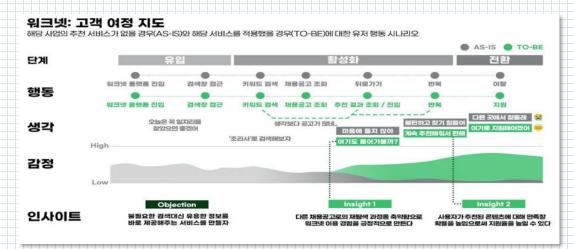
'찾으러가는' 일자리가 아닌 '가져다주는' 일자리로



워크넷에게 [좋은 일자리 추천]이란?

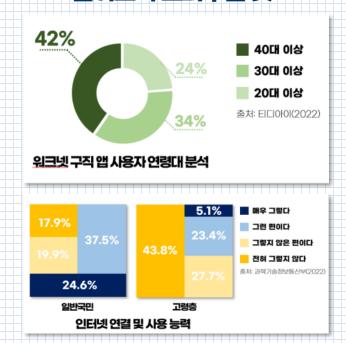
0

구직자들이 많이 클릭할 수 있는 매력적인 일자리를 제공하는 것



2

중장년층들의 디지털 정보격차를 줄이도록 도와주는 것







사업명: AI 언어모델을 이용한 유사 채용공고 추천 서비스 사업목표: 고성능의 언어모델 도입 및 추천 품질 향상

사업 핵심 기술: 자연어 처리, 파인 튜닝, 추천 알고리즘

개발제품명: W-BERT(Worknet-Bidirectional Encoder Representations Transformer)

핵심 데이터: 채용공고 API

성과 지표: 1인당 평균 방문 페이지 수, Hit-Rate(추천 성공률)

미케팅 컨셉:일 찾으러 다니는 건 그만! 고용아 잡(Job)줄래?

주요 EI겟층: 중장년 워크넷 이용자(5060)

USP: 채용 시장에 특화된 AI 언어 모델이 좋은 품질의 채용공고 추천을 진행함



002핵심기술

자연어 처리

복잡한 자연어 데이터를 학습 및 추천이 가능하도록 처리

W-BERT

Worknet-Bidirectional Encoder Representations from Transformers

파인튜닝

일반화된 모델에서 도메인 특화 모델로 추천 알고리즘

노출되면 안되는 콘텐츠들을 제어 및 유사한 문서 추천



002핵심 기술

학습: 현재 채용공고 문서가 어떤 직종을 가리키고 있는가?

학습데이터

정답데이터

차케어스 경기 용인시 수지구 주5일근무 관계없음 2110000.0 재가 요양보호사 그 외 보건 의료 종사자 재가 요양보호사 죽전동 요양보호사 금 근무 오전간 죽전동 한솔 노블빌리지 상생활 혼자함 식사챙겨드리기 청결유지 적극 연락 문의 자리 구함 으로 문자 남겨주세요 비정규직 시간선택제 학력무관 요양보호사1급 신분당선 판교 평일 오전 8시 30분 오전 11시 30분

보건 의료

마고워프 대전 서구 주5일근무 관계없음 2550000.0 총무 및일반 사무원 경영지원 사무 총무 및일반 사무원 주 마고워프 무역식품업 사무직원 채용 온라인 판매원 쇼핑몰택배 준비원 간단한회계업무 및 사무업무 엑셀 한글 워드 등 회사제품 출고 업무 송장출력 라벨지 작업 수량파악 출고현황 등 실업급여 부정수급을 위한 형식적인 채용신청은 관계자에게 통보됨을 알려드립니다 실제구직지원의사 있는 분들만 접수 비정규직 고졸 유성온천 평일 오전 9시 00분 오후 6시 00분 시간 40

경영 사무 금융 보험

재가, 요양, 청결, 의료, 식사··· 등은 보건 의료 분야에 속하는구나

경영지원, 엑셀, 쇼핑몰, 출력… 등은 경영 사무 금융 보험분야에 속하는구나

> 각 단어 및 토큰들이 워크넷이 지정한 직종 분야에 맞게 조정됨

감사원이 12일 공개한 '취업알선정보망 구축 및 관리 실태' 감사보고서에 따르면 AI 매칭 시스템의 입사지원율은 1차 사업(2020년 7월~2021년 7월) 9.45%, 2차 사업(2021년 7월~2021년 10월) 12.99%로 나타났다.

이는 당초 예측한 입사지원율인 59%, 70.2%에 크게 미치지 못하는 것이다.

나아가 같은 기간 워크넷의 기존 서비스인 <mark>'</mark>빅데이터 일자리 추천 서비스'의 입사지원율 15.22%, 18.36%<mark>보다도 낮았다.</mark>

감사원이 조회건수가 가장 많은 경기도 수원시를 표본 삼아 추천 구인정보 6천977건을 분석한 결과 구직자의 희망지역과 AI가 추천한 일자리가 일치하지 않는 경우가 1천20건으로 확인됐다.

심지어 희망지역을 수원으로 설정한 구직자에게 강원도 횡성군·태백시의 일자리를 추천<mark>하</mark>기도 했다.

워크넷의 기존 추천 서비스는 낮은 지원율과 다른 지역의 업체를 추천하는 등 0쉬운 점이 많이 남았던 서비스이다

반면 고용아 잡(Job)줄래 서비스는

1. 워크넷에서 정해준 옵션값이 아닌 구인자의 채용공고 전체를

INPUT으로 받기 때문에 하나의 채용공고에 대해 더 유사한 채용공고를 추천할 수 있다

- 2. 추천되어서 안될 콘텐츠를 노출시키지 않는 알고리즘을 적용했다(거리, 학력, 경력)
- 3. 당일에 올라오는 채용공고에 대한 추천이 이뤄지지 않는 점을 극복할 수 있다



003비교분석 및 성능 기존 서비스와의 차별점

[BEFORE] 추천 불가 콘텐츠

◎ 빅데이터 추천 정보 ☑ 연관된 추천정보가 없습니다.

당일 혹은 전일에 올라온 약 5,000건의 채용공고들에서는 추천이 전혀 되지 않는 상태

[AFTER] 추천 가능 콘텐츠



7만 개의 문서 기준 30분 이내로 추천이 가능하여 수시로 추천 아이템 업데이트 가능

또는 DB에서한건의채용공고등록시해당데이터에대한 추천만 갱신하는 방법도기능할것으로 추측



기존 서비스와의 차별점 추천되어선안될 콘텐츠 추천: 경력

[BEFORE]

필터링 미적용 콘텐츠



신입을 구하는 채용공고에 경력 필수 채용공고를 추천하는 것은 유사한 채용공고라고 할 수 없으며, 신입 지원자에게 불필요한 정보임

[AFTER] 필터링 적용 콘텐츠



신입을 구하는 채용공고에서 경력직 이상을 구하는 콘텐츠가 노출이 안됨



기존 서비스와의 차별점 추천되어선안될 콘텐츠 추천:학력

[BEFORE] 필터링 미적용 콘텐츠



고졸을 구하는 채용공고에 상위 학력을 요구하는 공고를 추천하는 것은 유사한 채용공고라고 할 수 없으며, 고졸 지원자에게 불필요한 정보임

【AFTER】 필터링 적용 콘텐츠



채용공고에서 제시하는 최소학력 이상을 요구하는 채용공고가 노출되지 않음 (중졸을구하는채용공고에서대졸을요구하는경우가없음)



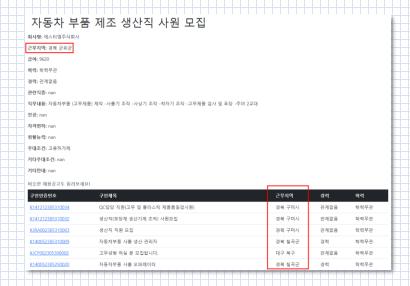
기존 서비스와의 차별점 추천되어선안될 콘텐츠 추천:거리

[BEFORE]

필터링 미적용 콘텐츠



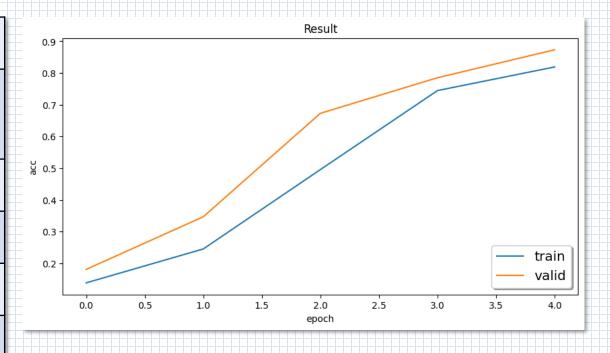
[AFTER] 필터링 적용 콘텐츠



충북 지역의 채용공고에 울산 지역의 채용공고를 추천하는 것은 유사한 채용공고라고 할 수 없으며, 충북 지원자에게 불필요한 정보임 수도권은 20km, 비수도권은 40km로 제한하여 지정된 거리 이상을 가지는 채용공고가 노출이 안됨

학습환경 및 파인튜닝 성능

학습환경	Google Colaboratory	
디바이스	GPU(V100)	
데이터 규모	약 6만 건	
학습 반복 횟수	5회	
배치 사이즈	32개	
소요 시간	56분	
파인튜닝 성능	87.5%	



> 시전 학습된 언어모델로 인해 빠르게 성능이 올라가는 모습 > 일반화된 모델(Kobert)에서 고용 시장 키워드를 효과적으로 학습한 도메인 특화 모델을 구축할 수 있음

추천 시간 약7만2천건기준

> 문서 벡터 추출 (2min)

W-BERT의 Word_embedding 레이어를 이용하여 문서를 표현하는 768차원의 벡터 생성

더이터처리(2min)

수집된데이터를추천기능한형태로정제하여 문장형식의데이터를생성

Facebook Al Similarity Search의 유사문서를 검색하는 기술을 사용하여 기존 코시인 유사도 기반의 검색 방식의 문제점(메모리, 속도)을 해결함

필터링 (20min)

특정 콘텐츠를 조회하는 유저들에게 **추천되어서 안될 콘텐츠를 필터링**함 Ex.서울 지역 일자리에서 부산 지역 일자리를 추천하는 상황 하지만 데이터 규모가 커질 수록 필터링을 수행하는 시간이 다소 지연되는 경향이 있음

완료

■총소요시간: 26분



제품 시연

질의응답

제 2회 고용노동부 공공데이터 활용 공모전

일 찾으러 다니는 건 그만! 고용아 잡(Job)줄래?

불필요한 검색을 줄여주는 채용공고 기반 추천서비스



