

2023-2

텍스트마이닝
팀 프로젝트 중간보고서

취업플랫폼 채용 정보 및 리뷰 분석 모델 제안

TEAM MEMBER

박상학(2023522122)
박세연(2023522114)
황성아(2023522119)

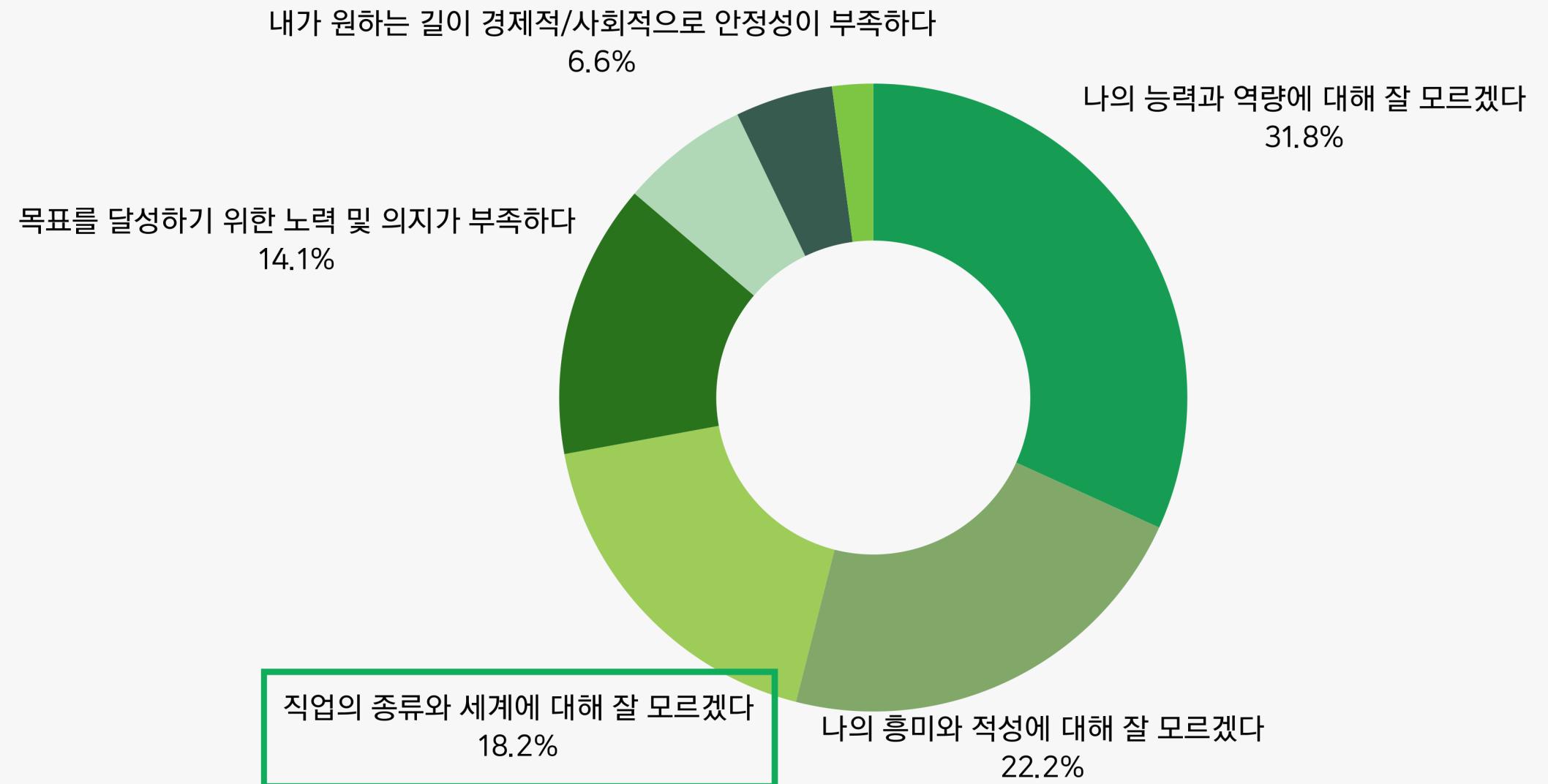
목 차

		Page
I	필요성	3
II	분석 주제	6
III	데이터 수집	7
IV	데이터 전처리	9
V	기업 리뷰 데이터 분석	11
VI	채용공고 데이터 분석	14
VII	활용방안	16
VIII	개선할 점	17
IX	TODO	18

| 필요성

진로 및 직업 탐색을 할 때, 직종별 어떤 역량을 준비해야 할지 막막한 상황이 많음

진로 결정/선택에 어려움을 경험하고 있다면, 그 이유가 무엇이라고 생각합니까?



- 진로 및 직업 탐색자는 특정 기업의 공고나 리뷰를 보는 것보다는, 해당 직종의 **공통적인 정보**를 더욱 필요로 함
- 관심 직종의 **전반적인 만족도**는 어떠한지, **공통적으로 요구하는 역량**에는 무엇이 있는지 등을 파악할 필요가 있음

| 필요성

직업에 대한 전반적인 역량이나 실제 맡게 되는 업무 정보를 얻기는 어려움

1) 흔히 접하는 정보와 실제 공고와의 불일치성

‘데이터 분석가 역량’에 대해 구글링한 결과, 엑셀 보고서를 잘 만드는 능력이 필요 요건 중 하나로 나옴

이러한 정보는 실제 요구 역량과 얼마나 일치할까에 대한 의문점

3. 데이터 분석가에게 필요한 것은?

지금까지의 설명을 통해 데이터 분석가의 업무 내용을 대략 이해하고 있을 거라고 생각합니다. 그러면 데이터 공부는 어디서부터 시작해야 할까요? 다음 항목에서는 데이터 애널리스트에게 필요한 기술을 살펴보았습니다.

| Excel- 엑셀 보고서

Excel은 표 계산 소프트웨어로서 널리 이용되면서 데이터 분석에도 중요한 역할을 하고 있습니다. Excel은 데이터 처리와 계산, 간단한 그래프 작성뿐 아니라 피벗테이블 만들기이나 Vlookup 등의 복잡한 함수를 포함한 고도의 기능도 구현할 수 있습니다. 데이터 애널리스트는 다른 데이터 분석이나 시각화 를 배우기 전에 Excel을 잘 사용해야 합니다.

구글링 결과

자격 요건

- 통계학, 컴퓨터공학, 정보보호학 등 관련 전공 학사 졸업자/졸업예정자 또는 그에 준하는 사 이버 위협 분석 관련 실무 경험/지식 (신입 가능)
- 데이터 분석을 위한 언어(Python, R 등)를 활용한 자동화 분석 경험
- 통계 및 데이터 분석 기술을 활용하여 인사이트를 도출해본 경험

실제 채용 공고의 자격요건

| 필요성

직업에 대한 전반적인 역량이나 실제 맡게 되는 업무 정보를 얻기는 어려움

2) 직종별 전반적인 공통 정보(요구 역량, 맡는 업무, 만족도 등)의 부재

취업 플랫폼에서는 현재 각 기업별 채용 공고와 리뷰는 확인할 수 있지만

직종별로 공통적으로 요구하는 역량이나, 직종별 만족도 등의 정보는 확인할 수 없음

The image shows a grid of job listing cards from different recruitment websites. The cards are arranged in two rows. The top row contains five cards: KEYENCE, S&I Corp., 배민, BKR, and ATRICS/AIRSMEDICAL. The bottom row contains four cards: sba 서울경제진흥원, Metanet, DR TECH, and daedong. Each card displays the company logo, job title, brief description, and some numerical data like 'D-3' or 'D-10'. Below the grid, there are two labels: '[사람인]' on the left and '[잡플래닛]' on the right, each pointing to its respective section of the grid.

[사람인]

[잡플래닛]

취업 플랫폼에서는 기업별로만 구분된 채용 공고, 리뷰를 확인할 수 있음

II 분석 주제

진로(직업) 탐색자를 위한 직무 비교, 직종별 주요 업무 및 자격 요건 요약, 만족도 정보 제공

1) 직종별 4가지 테마(업무와 삶의 균형, 복지, 회사 분위기, 연봉)에 대한 만족도 및 장단점 요약



2) 직종별 공통적인 키워드, 주요 업무, 자격 요건에 대한 요약

직무의 인기 키워드 Top 10				
keyword 키워드가 나온 문장 하나 정도 보여주기	keyword2 키워드가 나온 문장 하나 정도 보여주기	keyword3 키워드가 나온 문장 하나 정도 보여주기	keyword4 키워드가 나온 문장 하나 정도 보여주기	keyword5 키워드가 나온 문장 하나 정도 보여주기
keyword6 키워드가 나온 문장 하나 정도 보여주기	keyword7 키워드가 나온 문장 하나 정도 보여주기	keyword8 키워드가 나온 문장 하나 정도 보여주기	keyword9 키워드가 나온 문장 하나 정도 보여주기	keyword10 키워드가 나온 문장 하나 정도 보여주기

주요업무와 자격요건 및 우대사항

주요업무
백엔드 시스템 개발과 운영을 담당하는 Who We AreSaaS Platform 팀은 SSPSupplySide Platform의 백엔드 시스템 개발과 운영을 담당한다.

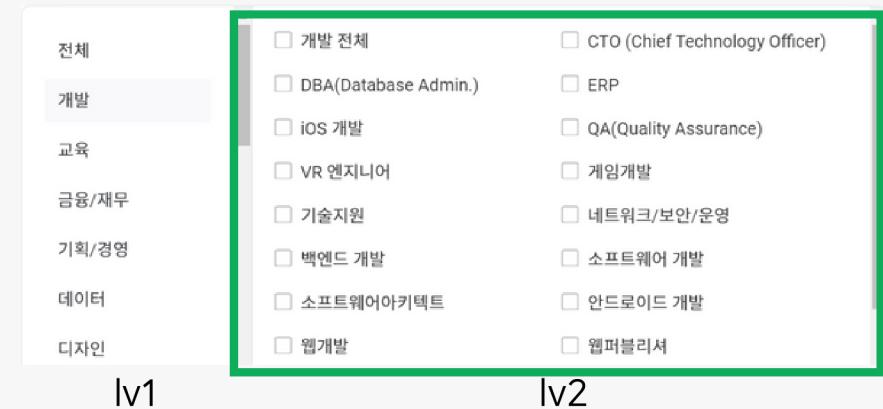
자격요건 및 우대사항
하나 이상의 서버 프로그래밍 언어에 능숙하여 테스트하기 좋은 코드를 작성하고 문제 해결 후 재발 방지를 위한 해법을 고민한다.

III 데이터 수집

직무 데스의 구성

1) 기업 채용 공고 데이터

- 잡플래닛의 현재 채용 공고가 올라온 기업에 대한 데이터 크롤링
 - 직무 하위 데프스(|v2)별 하나의 csv파일로 만들어 저장



[잡플래닛의 BI_엔지니어 직종 선택 시의 기업 공고]

총 702건	정확도순
 BI 엔지니어 캐롯손해보험(주) ★ 3.4 업무와 삶의 균형 : 3.9	<p>주요 업무</p> <ul style="list-style-type: none">• 캐롯 주행/계약 관련 지표 생성 및 제공해요.• Data-warehouse 및 Data mart 관련 스키마 설계 및 workflow 스케줄링을 해요.• Hadoop 기반 DW 구축 및 운영(w/Data engineer)을 해요.• Power BI 활용한 BI Dashboard 개발을 해요.
 Software Engineer (주)에어스메디컬 ★ 3.6 업무와 삶의 균형 : 4.6	<p>자격 요건</p> <ul style="list-style-type: none">• BI 및 ETL 관련 분야 경력 5년 이상이신 분이 필요해요.• DW 데이터 모델링 Skill이 있으신 분이 필요해요.
 기술연구소 기계장치 설계직 신입/경력 한국인닥타IRMWARE(주) ★ 2.7 승진 기회 및 가능성 : 3.0	<p>우대사항</p> <ul style="list-style-type: none">• BI 툴 활용 및 Hadoop에 대한 연동을 이해하신 분이면 좋아요.• BI 업무의 보다 효율적인 적용을 위한 경험을 보유하신 분이면 더 좋아요.• 시스템 구성상 전통적인 리포트툴로 적용하는 것보다 효율적인 오픈소스를 사용한 일부 시간성 리포트 생성 관리를 하시는 분이면 좋아요.
 DevOps 엔지니어 롯데헬스케어(주) ★ 4.0 복지 및 급여 : 4.0	<p>채용 절차</p> <ul style="list-style-type: none">• 해당 직무는 수시채용으로, 채용 완료 시 공고가 마감될 수 있어요.• 지원서를 허위로 기재할 경우 불합격 또는 입사가 취소될 수 있어요.• 국가보훈대상자 및 장애인은 관계법에 의거 우대해요.• 경력직의 경우 처우협의 단계에서 지원자의 정보 진위를 판단할 수 있는 증빙서류 제출이 요구될 수 있어요.
 커머스 백엔드 엔지니어 (개발 리더) 롯데헬스케어(주) ★ 4.0 복지 및 급여 : 4.0	



[BI 엔지니어_공고.csv]

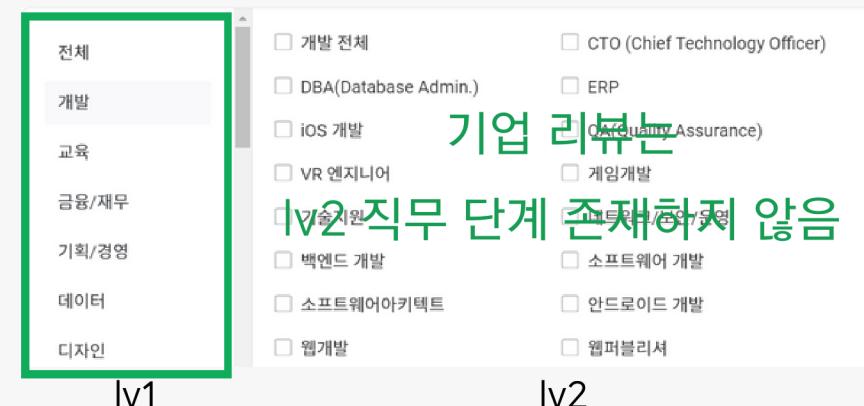
A	B	C	D	E	F	G	H	I	J	K	L	M	
1	회사명	고용 형태	스킬	기업 소개	주요 업무	자격 요건	우대 사항	채용 절차	복리후생	회사 위치	직무(복수)	직무_Jv1	직무_Jv2
				개조판에 보험은 한 화, SKT, 현대자동 차, 알토 스벤저스, 스틱인베 스트먼트 등 대형 투자사들 이 함께 합작하여 설립한 국 내 최초의 디지털 손 해보험사 입니다. 사물인터넷(IoT) 등 정보통신 기술(ICT) 를 결합하 어 더 간 BI	• 개조 수 행/계약 관련 지표 생성 및 제공해요. • Data- warehouse e 및 Data mart 관련 스키마 설 계 및 workflow 스케줄링 을 해요. • Hadoop 기반 DW 구축 및 운영 (w/Data engineer) 를 해요. • Power 를 결합하 BI 활용한 • BI 및 ETL 관련 분야 경력 의 보다 5년 이상 이신 분이 필요해요. • DW 데 이터 모델 링 Skill이 있으신 분 이 필요해 요.	• 비 글 글 용 및 Hadoop 에 대한 연동률 이 해하신 분 이면 좋아 요. • BI 및 ETL 관련 분야 경력 의 보다 5년 이상 이신 분이 필요해요. • DW 데 이터 모델 링 Skill이 있으신 분 이 필요해 요.	• 애정 식 무는 수시 채용으로, 시 공고가 해하신 분 있어요. • BI 업무 분이면 더 수 있어 면다면 더 좋아요. • 시스템 구성상 전 통적인 리 포트들로 적용하는 것보다 효 율적인 오 든소스를 • 비 글 글 무는 수시 채용으로, 시 공고가 마감될 수 있어요. • 지원서 를 하위로 의 보다 효율적인 적용을 위 한 경험을 보유하신 분이면 더 수 있어 면다면 더 좋아요. • 취소될 수 있어 면다면 더 좋아요. • 취소될 수 있어 면다면 더 좋아요. • 국가보 호법에 따라서 시스템 구성상 전 통적인 리 포트들로 적용하는 것보다 효 율적인 오 든소스를 • 애정 식 무는 수시 채용으로, 시 공고가 마감될 수 있어요. • 매년 복 지 포인트 200만원 • 설 차례 비 지급 • 체력단 련 휴가 5 일 • 시차출 퇴근제 • 자격증 취득 지원 • 무제한 온라인 교 육 지원 • 연차사 용 장려포 인트 35 만원 연차 숙 박비 지원	[정상]을 위한 휴 식] 재용 완료 시 공고가 지포인트 200만원 • 설 차례 비 지급 • 체력단 련 휴가 5 일 • 시차출 퇴근제 • 자격증 취득 지원 • 무제한 온라인 교 육 지원 • 연차사 용 장려포 인트 35 만원 연차 숙 박비 지원					
	캐롯손해노정규직	['Oracle', '']								서울 중구	['BI 엔지니어']	['BI 엔지니어']	

III 데이터 수집

직무 데스의 구성

2) 기업 리뷰 데이터

- 잡플래닛의 현재 채용 공고가 올라온 기업에 대한 데이터 크롤링
- 직무 상위 데스(lv1)별 하나의 csv파일로 만들어 저장



[잡플래닛 현재 채용 중인 기업 리뷰 페이지]

[기업리뷰.csv]

A	B	C	D	E	F	G	H	I	
1	직무	고용 현황	근무지	날짜	별점	요약	장점	단점	경영진에게 바라는 점
14	금융/재무	전직원	서울	2023. 01	3	"한화손보	연차를 아무때나 쓸 수 있고 젊은 조직이라 회사 생활은 재미있음	연봉협상이나 승진 등을 감추고 잘 알려주지 않아 투명성이 없음	젊은 조직답게 윗선에서 잘 해결해줬으면 하는데 피드백이 없음
15	개발	현직원	서울	2023. 01	3	"과대평가"	아직 대외적으로는 이미지가 좋은듯 하다. 신기술로 뭔가 인슈어테크 하는것 처럼 보임	속빈 강정이다. 좀더 내실을 다져야 하는데 눈앞에 실익에 급급함	인력 관리가 잘 안되는듯 하다. 리더 중에서도 노는 사람 많다
16	개발	현직원	서울	2022. 12	3	"워라밸 좋 및 체력단련 휴가지 하철역에서 바로 회사로 출퇴근 할 수 있음"	눈치 안보고 연차 쓸수 있음 넉넉한 복지포인트	회사 공지가 직원들에게 제대로 공유되지 않음 누구는 알고 누구는 모르고 최근에 새로운 대표님 오시면서 많이	마케팅에 집중하기보다 고객들이 보다 신뢰할 수 있는 퍼마일 모르고 최근에 새로운 대표님 오시면서 많이 방향을 확실하게 잡
17	마케팅/시	전직원	서울	2022. 10	1	"겉만 번지 개념 자체는 나쁘지 않아서 많이 팔리기는 했다. 그러나..."	모회사가 대기업이라는 정도? 퍼마일 보험 상품	농구단 문제로 뉴스 탈 만큼 자금 사정 좋지 않다. 정말 근본적인 문제는 플러그 리스크	하루 빨리 관종형 무능력자 걸러내지 못하면 미래는 없어 보인다.



IV 데이터 전처리

- 중복 열 삭제, 데이터 타입 변환, 주요 컬럼 제외 결측값 존재 행 삭제 등 기본적인 전처리 완료
- 컬럼과 일치하지 않은 내용이 다수 존재하므로, 해당 컬럼 관련 주제가 아닌 다른 내용은 삭제

고 용 형 태	회사 명	스킬	기업 소개	주요 업무	자격 요건	우대 사항	채용 절차	복리후생	회사위치	직무 _lv1	직무 _lv2
2	잡플래닛	정규직	[데이터분석, BI, python, SQL, Tableau, POWER BI]	더 나은 선택은 없을까? 지금 이 순간에도 잡플래닛을 방문하는 월 300만명...	[조직소개] [조직조건] [우대사항] [자격 조건] [Data Visualiz...]	[전형 프로세스] [서류전형] -> 1차 실무 인터뷰 -> 2차 협업부서 인터...	- 분 단위 시차 출퇴근제 (오전 7시 ~ 오전 10시 사이 자율 출퇴근) - 분 단위 시차 출퇴근제 (오전 7시 ~ 오전 10시 사이 자율 출퇴근)	서울특별시 강남구 테헤란로 509 엔씨타워 10층	[데이터]	[BI 엔지니어]	

‘주요 업무’ 컬럼에 기업 소개 관련 내용이 함께 포함되어 있음



고 용 형 태	회사 명	스킬	기업 소개	주요 업무	자격 요건	우대 사항	채용 절차	복리후생	회사위치	직무 _lv1	직무 _lv2	자격요건 및 우대사항
2	잡플래닛	정규직	[데이터분석, BI, python, SQL, Tableau, POWER BI]	더 나은 선택은 없을까? 지금 이 순간에도 잡플래닛을 방문하는 월 300만명...	[전사 Business Intelligence : Dashboard, R...]	[자격 조건] [우대사항] [Data Visualiz...]	[전형 프로세스] [서류전형] -> 1차 실무 인터뷰 -> 2차 협업부서 인터...	- 분 단위 시차 출퇴근제 (오전 7시 ~ 오전 10시 사이 자율 출퇴근)	서울특별시 강남구 테헤란로 509 엔씨타워 10층	[데이터]	[BI 엔지니어]	[자격 조건] [SQL 상급, 파이썬 증급] [Data Visualiz...]

IV 데이터 전처리

- 분석에 사용할 컬럼에 대해 문자열 정제

고 용 형 태	회사 명	스킬	기업 소개	주요 업무	자격 요건	우대 사항	채용 절차	복리후생	회사 위치	직무 _lv1	직무 _lv2	자격요건 및 우 대사항
2	잡플 래닛	[데이터분석, BI, python, SQL, Tableau, POWER BI]	더 나은 선택 은 없을까? \n 지금 이 순간에도 잡 플래닛을방 문하는 월 300만명...	[자격 조건] 전사 Business Intelligence : Dashboard, R...	[우대 사항] SQL 상급, 파이 썬 중급 Data Visualiz...	[전형 프로세 스] 서류전 형 -> 1차 실무 인터뷰 -> 2차 협업부서 인 터...	- 분 단위 시차 출퇴근제 (오전 7시~ 오전 10 시 사이 자율 출퇴근)\r\n ...	서울특별 시 강남구 테헤란로 509 엔씨 타워 I, 10 층	[데이 터]	[BI 엔 지니 어]	[자격 조건] SQL 상급, 파이 썬 중급 Data Visualiz...	

\Wr, \Wn, 이모지 등이 포함된 문자열



고 용 형 태	회사 명	스킬	기업 소개	주요 업무	자격 요건	우대 사항	채용 절차	복리후생	회사위 치	직무 _lv1	직무 _lv2	자격요건 및 우 대사항	index
2	잡 플 래 닛	[데이터분 석, BI, python, SQL, Tableau, POWER BI]	더 나은 선택은 없을까?\n 지금 이 순간에도 잡 플래닛을 방문하 는 월 300만명...	[자격 조 건]\n 전사 Business Intelligence Dashboard Report 개발...	[우대 사항]\n SQL 상급, 파이썬 중급\n Data Visualiz...	[전형 프로 세스]\n 서류전 형 -> 1차 실무 인터뷰 -> 2차 협업부서 인 터...	- 분 단위 시차 출퇴 근제 (오전 7시~ 오전 10시 사이 자율 출퇴 근)\r\n ...	서울특 별시 강 남구 테 헤란로 509 엔 씨타워 I, 10층	[데이 터]	[BI 엔지 니어]	자격 조건 SQL 상급 파이썬 중 급 Data Visualiz ation 소 프트웨어Ta...	2	

▽ 기업 리뷰 데이터 분석

1) 테마별 감성분석을 위해 형태소 분석을 통해 리뷰 중 많이 언급되는 키워드를 추출해 테마 생성

{'회사': 25773,
'수': 21920,
'업무': 21750,
'사람': 21315,
'일': 17547,
'직원': 15281,
'분위기': 14210
'복지': 13974,
'연차': 13341,
'것': 11878,
'팀': 11674,
'야근': 11318,
'함': 10979,
'연봉': 10643,
'눈치': 7846,
'사용': 7766,
'문화': 7066,
'기업': 7014,
'생각': 6963,
'근무': 6308,
'체계': 5958,
'편': 5945,
'시간': 5629,
'등': 5577,
'장점': 5485,
'지원': 5373,
'때': 5347,
'부서': 5335,
'가능': 5109,
'비': 4819,
'퇴근': 4680,



형태소 분석을 통한 리뷰 데이터
Top100 명사 키워드 추출

테마

업무와 삶의 균형

해당 키워드

업무, 일, 야근, 근무, 시간, 퇴근, 자유, 경험, 성장, 능력, 기회, 강도

회사 분위기

직원, 분위기, 팀, 눈치, 문화, 체계, 수평, 환경, 조직, 제도

복지

복지, 연차, 휴가, 식대

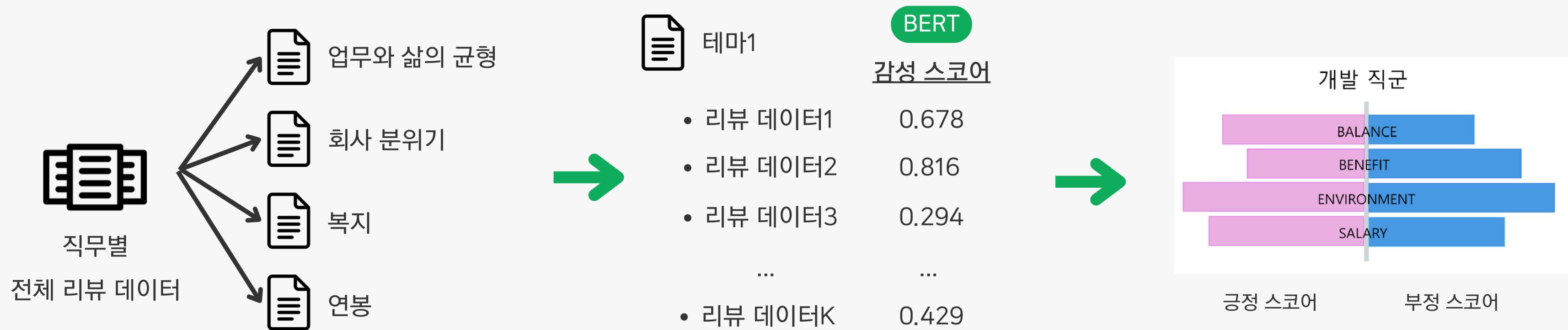
연봉

연봉, 급여, 월급, 돈

추출한 키워드 단어를 비슷한 주제끼리 묶어 4가지의 테마 생성

▽ 기업 리뷰 데이터 분석

2) 직무(lv1)의 테마별 감성 스코어 계산



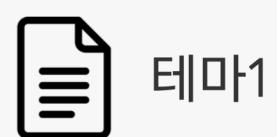
직무의 전체 리뷰 문장을 4가지 테마에 맞게 분류

BERT pretrained model을 이용해
테마별 리뷰 데이터마다의 감성 스코어 추출

한 테마당 긍정의 정도(감성 스코어 값 0.5 이상)와 부정의 정도(0.5미만)의 평균을 구
해 최종 긍정 스코어, 부정 스코어 계산

▽ 기업 리뷰 데이터 분석

3) 직무(lv1)의 테마별 리뷰 데이터 요약



- 리뷰 데이터1
- 리뷰 데이터2
- 리뷰 데이터3
- 리뷰 데이터K



테마1 모든 리뷰 데이터



T5

e.g. 영업제휴 - ‘업무와 삶의 균형’ 부정 리뷰 요약본

근무시간이 길며 업무의 단순화로 일을 하는 실무자로선 매너리즘
에 빠지게 되는 경우가 많고 업무 방향이 자주 변경됨 부바부지만
야근이 있음

T5 pretrained model을 이용해 테마별 리뷰 요약글 생성

BART

e.g. 영업제휴 - ‘업무와 삶의 균형’ 부정 리뷰 요약본

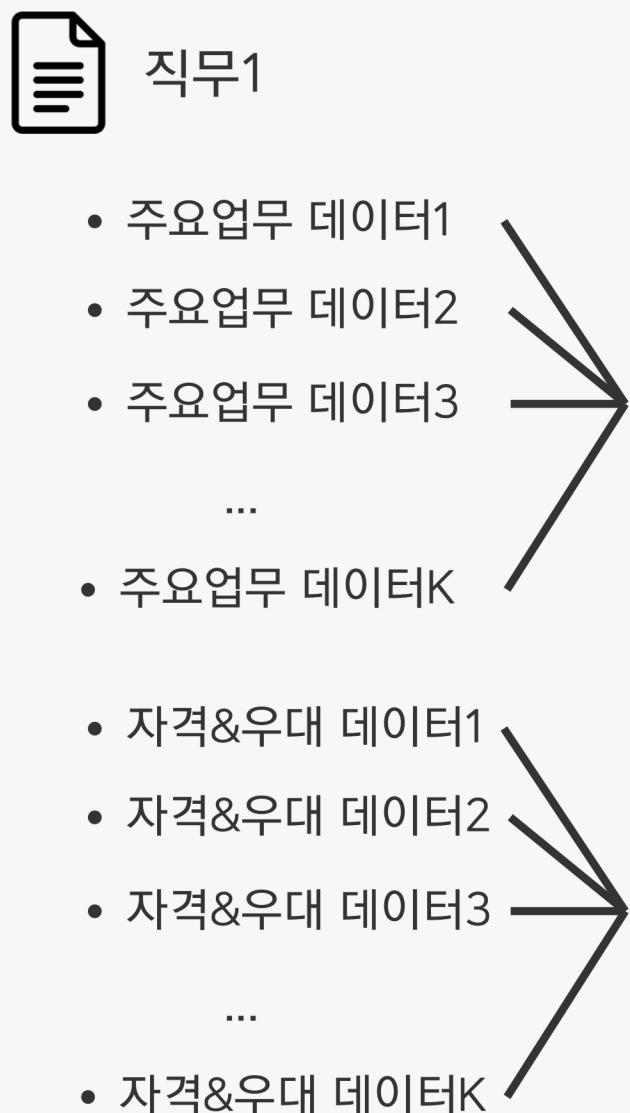
동종 업계 대비 연봉 수준 낮은 편근데 사내 문화도 후진적IT 회사지만
재택 근무는 찾아볼 수도 없고 임원 얼굴 외워서 인사하라고 지시 내려옴
니다

BART pretrained model을 이용해 테마별 리뷰 요약글 생성

직무의 테마별 리뷰 데이터를 하나의 str 타입으로 합치기

VI 채용공고 데이터 분석

1) 직무(lv2)의 채용공고(주요업무, 자격요건및우대사항)의 키워드(한글, 영어) Top10 추출



직무1 모든 주요업무



okt
nltk

한글, 영어 명사 형태소
종합 Top 25 추출



okt
nltk

한글, 영어 명사 형태소
종합 Top 25 추출

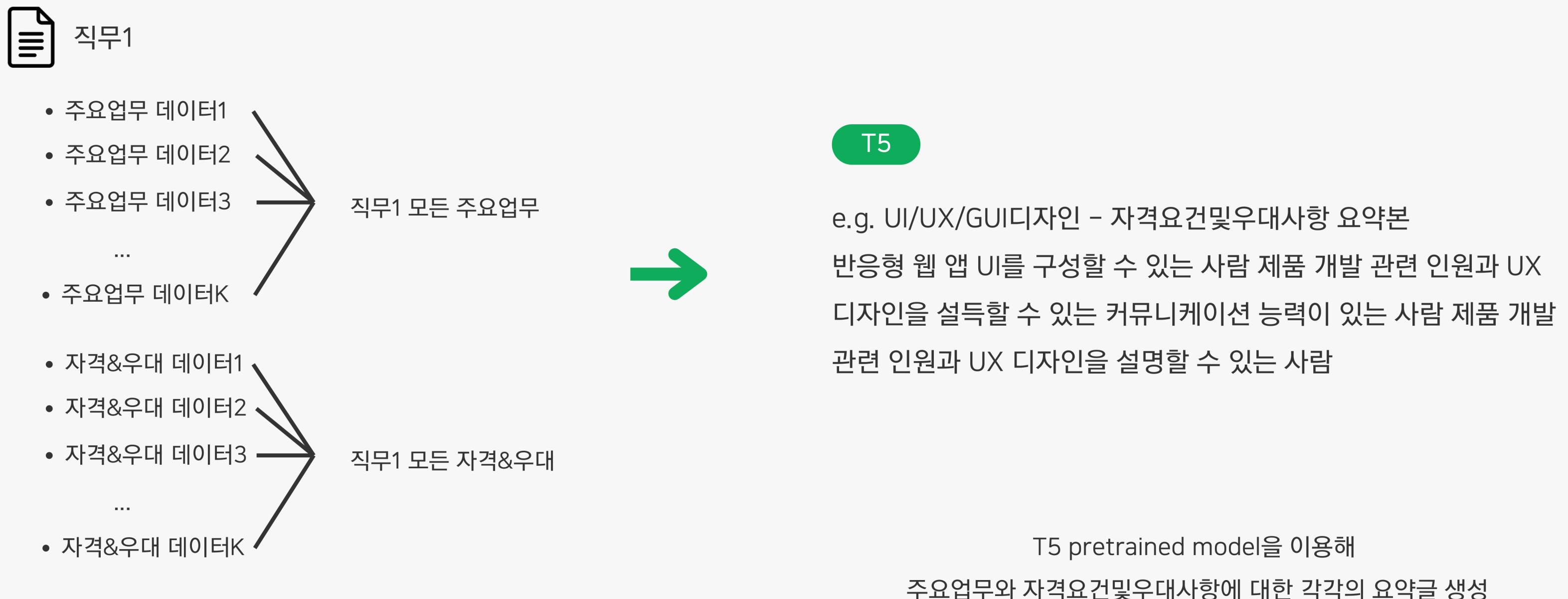
okt, nltk 모델을 이용해 형태소 분석

e.g. 웹개발 - 자격요건및우대사항
Top25 키워드

운영: 173
웹: 141
설계: 136
기술: 94
프로젝트: 89
환경: 83
시스템: 77
협업: 76
React: 74
문제: 70
커뮤니케이션: 69
구축: 68
역량: 64
AWS: 62
코드: 52
API: 52
지식: 49
언어: 47
이용: 47
Spring: 47
관리: 46
백엔드: 44
테스트: 44
Java: 43
프론트엔드: 42
Git: 41
프레임워크: 39
배포: 39
Python: 39
서버: 37

VI 채용공고 데이터 분석

2) 직무(lv2)의 채용공고(주요업무, 자격요건및우대사항) 데이터 요약



VII 활용 방안

Contribution

- 실제 직종별 현업 종사자의 리뷰 및 실제 채용 공고에 대한 요약 및 정리를 제공함으로써 진로탐색자에게 보다 신뢰도 높은 정보 제공

나아가, 분석 결과에 대해 사용자가 보기 쉽게 웹으로 제작(진행 중)

취업플랫폼 채용 정보 및 리뷰 분석 직무선택파악 직무분석 Search Search

직무 성향

직무의 테마별 감성분석

직무	복지	분위기	연봉
업무와 삶의 균형	40/100	41/100	40/100
기술/경영	복지	분위기	연봉
마케팅/시장조사	41/100	40/100	41/100
미디어/홍보	복지	분위기	연봉
생산관리/품질관리	복지	분위기	연봉
서비스/고객지원	복지	분위기	연봉
연구개발	복지	분위기	연봉
유통/무역	복지	분위기	연봉
전문직	복지	분위기	연봉

직무 분석

직무의 테마별 감성분석

직무 분석

취업플랫폼 채용 정보 및 리뷰 분석 직무선택파악 직무분석 Search Search

직무 분석

직무의 인기 키워드 Top 10

keyword	keyword2	keyword3	keyword4	keyword5
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				
키워드가 나온 문장 하나 정도 보여주기				

직무의 인기 키워드 Top 10

직무 분석

직무 분석

VIII 개선할 점

1) 요약 태스크에 대해 전반적으로 낮은 성능

(회사명 등이 포함되는 등 특정 기업에만 해당되는 내용 위주로 요약됨, 전처리의 문제인지?)

```
review_summarization > result > 개발_neg_balance_smr.txt
1 대기업의 단점 과 스타트업 단점이 있는곳 부서간 이해관계나 금융업계의 보수적인 업무처리 방식등이 단점으로 생각된다.
```

2) 채용 공고-주요업무 요약 태스크

구체적인 정보를 얻는 게 목표인데, 한 줄 요약으로는 부족한 듯

```
review_sentiment.ipynb M 패션설유디자인_work.txt X
gongo_summarization > result > 패션설유디자인_work.txt
1 입사하면 담당할 업무는 시즌 그래픽 및 아트워크 개발 시장조사 및 트렌드 조사 분석
컨펌 준비 및 품평 업무 업체 핸들링 및 유관부서 협업 담당 디자인 핸들링 및 제품 QC 진행이다.
```

3) 감정분석 성능 비교

감정분석에서 koBERT를 일단 사용하여 점수를 추출했는데, 사전 학습된 모델을 그냥 사용하다 보니 성능 비교를 어떻게 하면 좋을지

4) 리뷰 데이터의 테마(워라밸, 회사분위기, 연봉, 복지) 4가지인데, 테마 개수를 늘릴 필요가 있을지

5) 리뷰 데이터는 채용공고 데이터와 다르게 상위 레벨(lv1)으로만 데이터가 존재해, 세부 직무별 감성분석을 하진 못함

6) 활용 방안이 구체적이지 못함

(이러한 분석으로 비즈니스적으로 어떤 활용이 가능할지, 어떤 방식으로 더 다양한 분석을 진행해야 하는지)

IX TODO

- 직무 비교 대시보드 만들 예정
- 요약에서 기업 세부적인 내용이 나오지 않게끔 추가 전처리
- 사전학습 모델 돌릴 때, 추가적인 하이퍼 파라미터 튜닝
- Bart(추출요약)과 T5(생성요약)모델 비교
- 웹 페이지 남은 부분 개발

감사합니다