

팀장: 박세연(2023522114)

박상학(2023522122)

황성아(2023522119)

# 취업플랫폼 채용 정보 및 리뷰 분석 모델 제안



# 취준생의 니즈

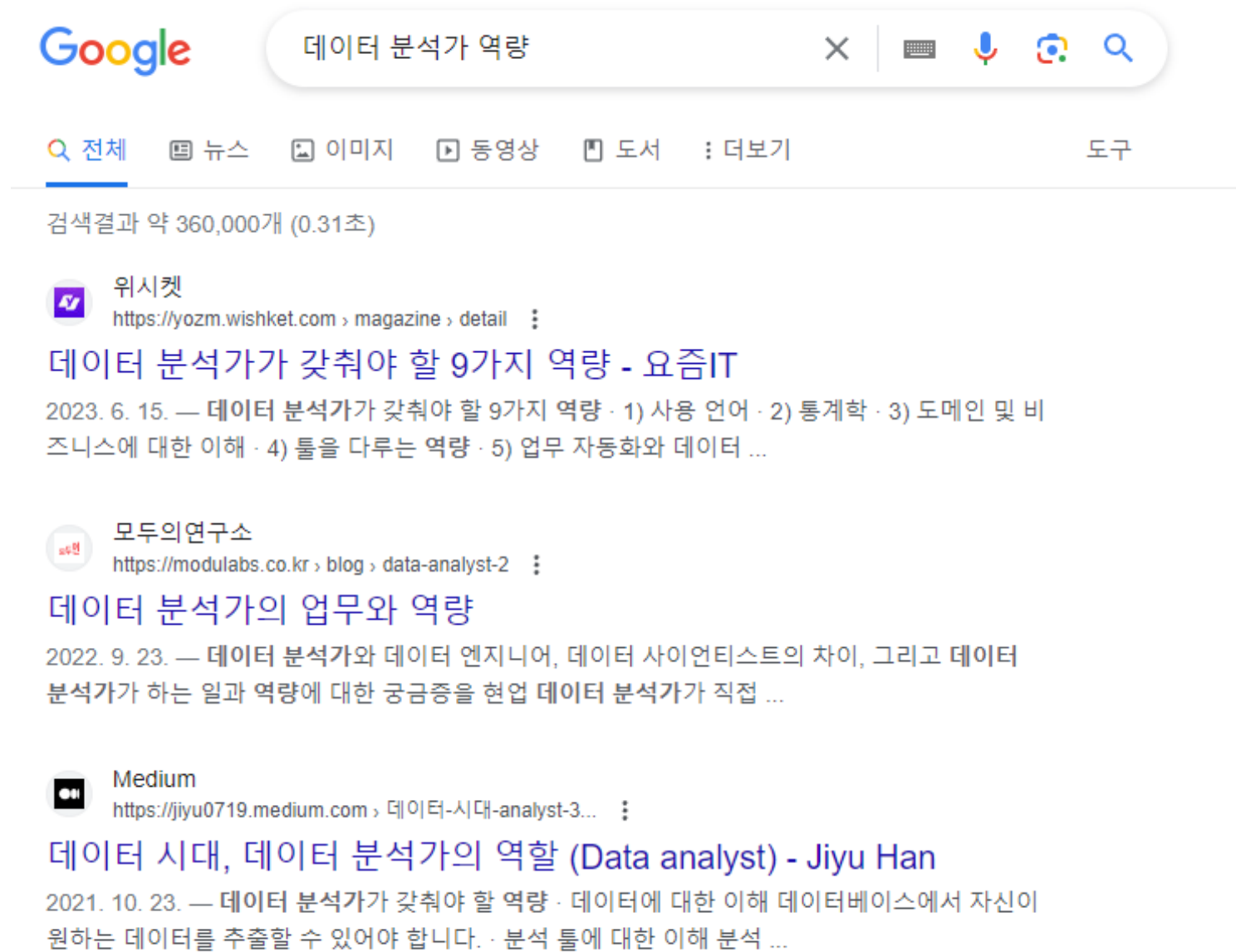
직종별 어떤 역량을 준비해야 할지 막막할 때가 존재

특정 직종에 관심있어 관련 스펙을 이제 막 쌓아나가려는 취준생들은  
원하는 직장이 하나로 명확히 정해지진 않은 상태,  
그러나 그 직종에서 공통적으로 요구하는 역량을 키워나가고 싶은 욕구

혹은, 내가 지금까지 쌓아온 역량은 어떤 직종과 잘 맞을지,  
그 직종의 만족도는 전반적으로 어떠한지 궁금할 때

# 현재의 한계점

구글링하면 필요한 직무 역량을 써놓은 글들이 많지만, 얼마나 믿을만한 정보일까?



## 3.데이터 분석가에게 필요한 것은?

지금까지의 설명을 통해 데이터 분석가의 업무 내용을 대략 이해하고 있을 거라고 생각합니다.그러면 데이터 공부는 어디서 부터 시작해야할까요? 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 만들기**이나 Vlookup 등의 복잡한 함수를 포함한 고도의 기능도 구현할 수 있습니다.데이터 애널리스트는 다른 데이터 분석이나 시각화 툴을 배우기 전에 Excel을 잘 사용해야 합니다.

### SQL : SQL 개발자를 위해

위키백과 정의에 따르면, SQL( Structured Query Language, 구조화 질의어)는 관계형 데이터베이스 관리 시스템(RDBMS)의 데이터를 관리하기 위해 설계된 특수 목적의 프로그래밍 언어입니다. RDBMS가

-> 데이터 분석가가 되기 위해서는 엑셀을 잘 사용하는 능력이 가장 필요한 요건 중 하나일까?

# 현재의 한계점

현재 직종별 채용중인 기업의 채용정보를 볼 수는 있지만,  
직종별 공통적으로 필요한 역량이나, 전반적으로 맡게 되는 업무 등에 대한 정보를 얻기는 어려움

데이터 분석가 ▼

경력 ▼


근무 지역 ▼

평점 ▼


고용 형태 ▼

학력 ▼

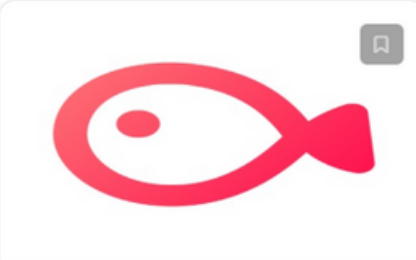
초기화




**데이터분석가 (주니어)**  
3년 이상, SQL, 커뮤니케이션  
(주)고픽커뮤니티 ★ 0.0 (0)  
취업 축하금: 200만원




**데이터 분석가**  
3년 이상, 기획, 데이터분석  
(주)수퍼트리 ★ 2.9 (19)  
취업 축하금: 200만원




**데이터사이언티스트 (경력)**  
2년 이상, 마케팅, 시장조사, python, SQL, 머신러닝  
(주)비모소프트 ★ 2.7 (22)  
취업 축하금: 200만원



**데이터 분석가 (경력, 리더)**  
5년 이상, 데이터분석, python, SQL  
(주)에이에프아이 ★ 4.3 (6)  
취업 축하금: 200만원



**GA AA 분석가**  
1년 이상, 데이터분석, 데이터 분석, google analy...  
콘센트릭스서비... ★ 2.5 (662)  
취업 축하금: 200만원



**데이터 추출 담당자(1년 계...)**  
경력무관, python, SQL, 데이터, 추출  
(주)코인원 ★ 3.4 (96)  
취업 축하금: 200만원

## 주요 업무

다양한 R&D분야의 HPC 고성능 컴퓨팅 플랫폼 구축 업무  
HPC 상에서 구동되는 응용 개발 환경을 구성하고 서비스 환경으로 전환하는 고급 엔지니어 업무

- Linux-HPC 클러스터 시스템 설치 및 운영
- 연구개발 분야 응용 개발 환경 설치 및 운영
- VMWare/XenServer 가상화 솔루션 기반 대규모 VDI 인프라/플랫폼 구성 사업 기술 지원
- 대기업/대규모 가상화 시스템 아키텍처 설계/제안/프로젝트관리
- AWS EC2, Naver Cloud Platform 등 클라우드 인프라 및 플랫폼 구성 사업 지원

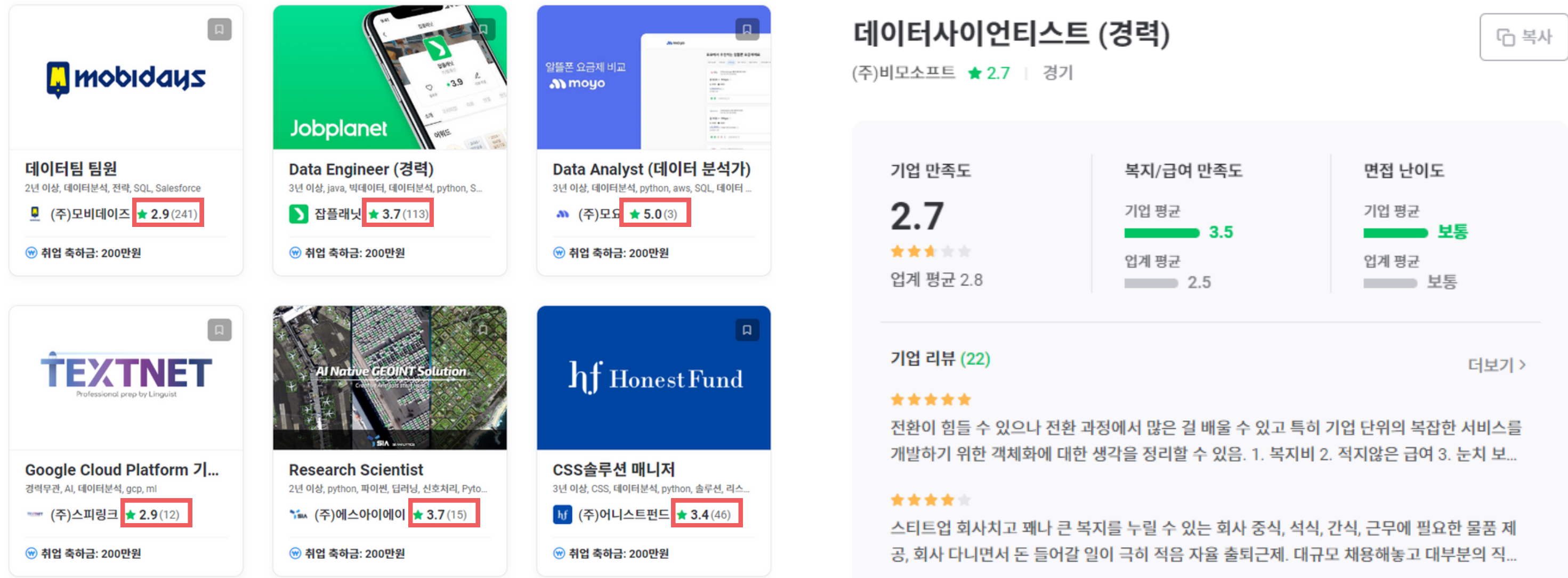
## ● 이런 분과 함께하고 싶어요

- 데이터 분석 관련 실무 경력 2년 이상이신 분
- 데이터 분석, 머신러닝 또는 데이터 기반 마케팅에 대한 전반적인 지식을 보유하신 분
- 데이터 추출, 전처리, 분석, 인사이트 도출까지 데이터 분석 전과정을 경험해보신 분
- SQL을 자유자재로 활용하여 Raw Data 처리 및 분석이 가능하신 분
- 분석 결과를 통한 결과 도출 및 논리적인 개선사항을 제시할 수 있는 분
- 데이터 시각화(Tableau, Amplitude 등)에 대한 경험을 보유하신 분
- 기획, 마케팅, 개발 등 다양한 직군과 원활한 커뮤니케이션 및 협업이 가능하신 분

# 현재의 한계점

한 기업에 대한 만족도는 알 수 있으나 해당 직종에 대한 만족도는 알아내기 어려움

특정 직종에 관심이 생긴 취준생에게는 직종을 결정하는 중요 요소일 수 있음



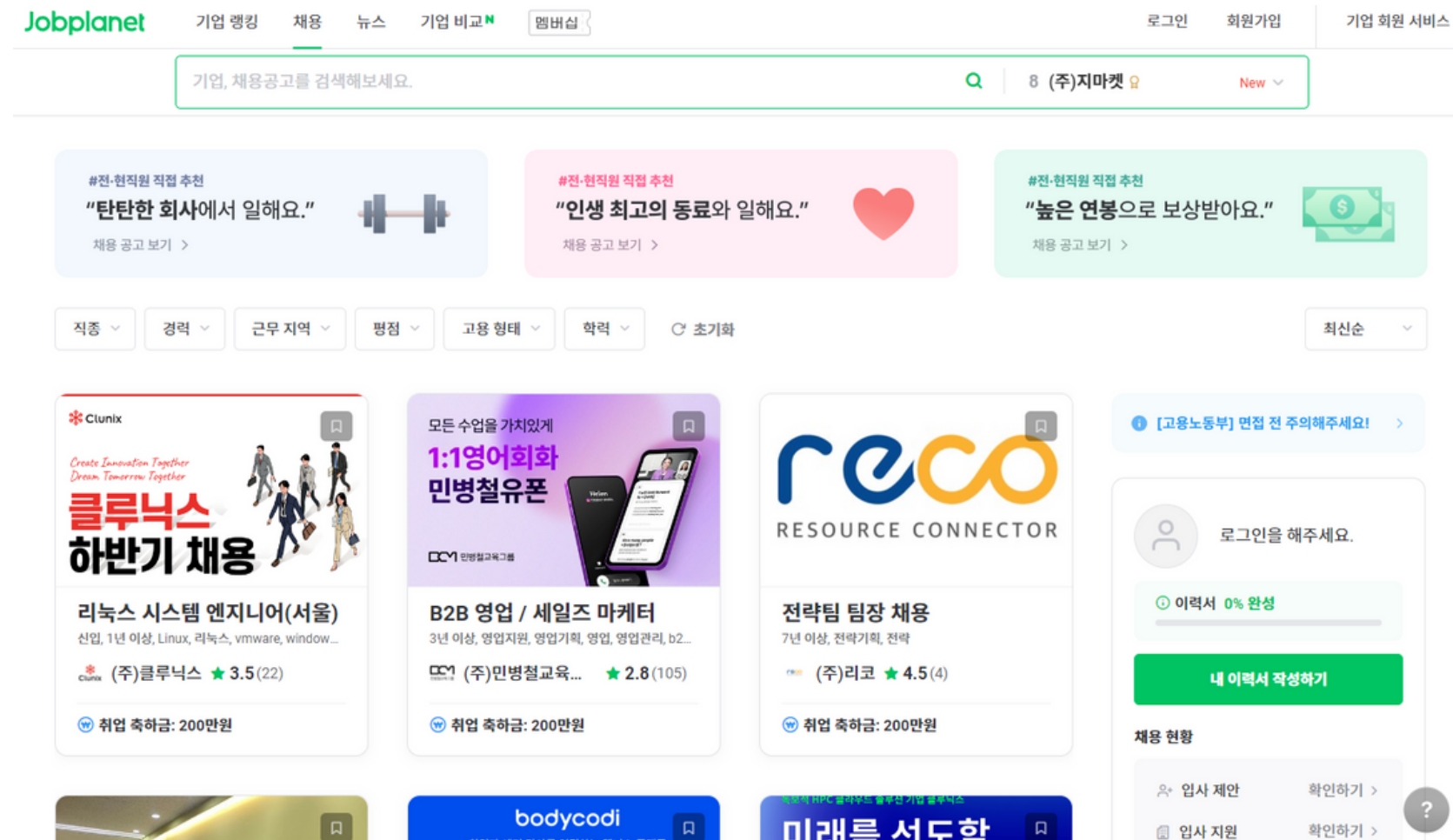


# 데이터 선정

## 잡플래닛

국내 취업플랫폼 중 5순위 안에 드는 인기 있는 플랫폼

다른 취업플랫폼과 달리 기업 리뷰와 채용 정보를 한 번에 볼 수 있는 장점



# 분석 주제

- 1) 직종별 특성에 대한 만족도를 분석
  - 2) 직종별 주요 업무, 자격 요건에 대해 요약 정리한 정보 제공
- 실제 현업 종사자의 리뷰와 실제 채용 공고를 통해 보다 신뢰도 높은 정보 제공

직종 ^

경력 v

근무 지역 v

평점 v

고용 형태 v

학력 v

초기화

교육

금융/재무

기획/경영

데이터

디자인

마케팅/시장조사

미디어/홍보

☐ 데이터 전체

☐ 데이터 분석가

☐ 데이터 엔지니어

☐ 빅데이터 엔지니어

☐ BI 엔지니어

☐ 데이터 사이언티스트

☐ 머신러닝 엔지니어

직종 분류

기획/경영 | 현직원 | 서울

기업 추천 리뷰

★★★★★

"코시국에 직원들에게 제공하는 복지문화 Good"

승진 기회 및 가능성

복지 및 급여

업무와 삶의 균형

사내문화

경영진

장점

주니어에서 시니어로 넘어오는 경력자들에게 추천

경험해온 내용들을 체계적으로 펼칠수가 있다.

7시부터 10시까지 분단위로 출근할 수 가 있음

연차가 많아서 번아웃될뻔? 하다가도 예방할 수 가 있다ㅋㅋㅋㅋ여름휴가 돌봄휴가 ...유급병가까지.

ㅋㅋ

단점

재택근무가 장기화되면서 회사에서 소속감을 찾을 수 있는 활동이 거의없다.

경력자에게는 추천하는 회사이지만, 신입 주니어 레벨에서는 다소 높은 업무강도가 있을 수 있다.

경영진에 바라는 점

수익이 늘어나는 만큼 연봉도 올려주실꺼죠?? ㅎㅎㅎㅎㅎㅎ

이 기업은 1년 후 성장하고 있을 것이다.

이 기업을 추천 합니다!

도움이 돼요 24

페이스북에 공유

신고하기

기업에 대한 개인 리뷰

# 분석 주제

## 1) 직종별 만족도 분석

: 잡플래닛의 직종별 기업 리뷰를 크롤링해 감성분석을 진행하는 방식



Selenium과 BeautifulSoup을 이용한 기업의 직종별 리뷰 크롤링

### KoBERT

Korean BERT (Bidirectional Encoder Representations from Transformers)



KoBERT는 기존 BERT의 한국어 성능 한계를 극복하기 위해 개발되었다. 위키피디아나 뉴스 등에서 수집한 수백만 개의 한국어 문장으로 이루어진 대규모말뭉치(corpus)를 학습하였으며, 한국어의 불규칙한 언어 변화의 특성을 반영하기 위해 데이터 기반 토큰화(Tokenization) 기법을 적용하여 기존 대비 27%의 토큰만으로 2.6% 이상의 성능 향상을 이끌어 냈다.

대량의 데이터를 빠른시간에 학습하기 위해 링 리듀스(ring-reduce) 기반 분산 학습 기술을 사용하여, 십억 개 이상의 문장을 다수의 머신에서 빠르게 학습한다. 더불어, 파이토치(PyTorch), 텐서플로우(TensorFlow), ONNX, MXNet을 포함한 다양한 딥러닝 API를 지원함으로써, 많은 분야에서 언어 이해 서비스 확산에 기여하고 있다.

크롤링 데이터 전처리 후 감성분석 모델을 이용해 리뷰에 대한 감성분석

### 데이터

복지



급여



업무와 삶의 균형



사내문화/분위기



성장 가능성



감성score값을 이용해 직종의 특성별 만족도(급여, 사내문화/분위기 등) 나타내기



# 분석 주제

## 2) 직종별 업무/역량 분석 및 요약

: 잡플래닛의 직종별 현재 채용 중인 기업의 주요 업무와 자격 요건 등의 텍스트 데이터를 크롤링해 텍스트 요약

### 주요 업무

[함께하시면 맡게 될 업무입니다]

- 머신러닝 알고리즘 기술 연구 및 개발
- 시계열 딥러닝 알고리즘 기술 연구 및 개발
- 제조업 불량예측 및 Anomaly detection에 적용 가능한 알고리즘 개발
- 알고리즘 고도화

### 자격 요건

[이런 분이 필요해요!]

- Python을 활용하여 머신러닝/딥러닝 프로젝트 (Prediction, Anomaly detection, Classification) 경험자
- Tensorflow 또는 Pytorch 사용 경험
- Pandas, Matplotlib등의 패키지를 활용한 데이터 정제, 전처리, 분석, 모델링, 시각화 경험을 보유하신 분
- 최신 기술에 관심이 많은 분
- 주도적으로 일을 만들고, 책임감 있게 리드하시는 분
- 다른 직군과 함께 문제를 정의하고 해결책을 찾을 수 있는 분

[업무에 꼭 필요한 기술 스택입니다]

- Tensorflow
- Pytorch

### 한국어 이해생성 언어모델(ET5)

» 학습 모델 및 데이터 신청 바로가기

» 학습 데이터 신청하기

» 학습 데이터 다운로드

» 한국어 이해생성 언어모델(ET5)

과학기술정보통신부와 ITP의 혁신성장동력 프로젝트로 추진 중인 **엑소브레인 사업에서 한국어 이해 및 생성이 가능한 언어모델**을 공개합니다.

ETRI 엑소브레인 연구진이 배포하는 한국어 이해생성 언어모델(ET5)은 **한국어를 이해하고 실시간으로 출력 텍스트를 생성**해 낼 수 있어서, 한국어분석, 기계독해, 문장의미 분석과 같은 언어 이해 유형의 태스크뿐 아니라 문서요약, 질문생성 등과 같은 다양한 언어 생성 유형의 태스크에도 활용이 가능합니다.

한국어 이해생성 언어모델(ET5)은 대용량 원시 텍스트로부터 **빈칸 단어열 맞추기(T5 학습 유형)**와 **다음 단어 맞추기(GPT 학습 유형)**를 동시에 사전학습(pre-train)하여 **언어이해와 언어생성 능력을 향상**하였습니다. 대표적인 한국어 처리 태스크 5종(기계독해, 요약, 단락 순위화, 형태소 분석, 문장유사도 추론) 대상 평가 결과 한국어 최고 수준의 성능을 보였습니다. 구글에서 배포한 모델에 비해 기계독해 1.40%, 형태소 분석 1.43%, 문장유사도 추론 7.12% 우수한 성능으로 평가되었습니다.

### 데이터

### 주요 업무

머신러닝 및 딥러닝의 알고리즘 기술을 연구하고 개발하는 업무를 ...

### 자격 요건

업무에 필요한 필수 기술 스택으로 는 Tensorflow, Pytorch이 있으며, ...

기업의 채용 공고 텍스트 데이터 크롤링

기업의 주요 업무와 자격요건에 대한 정보 요약 정리

직종별 요약된 주요 업무,  
자격요건 정보 제공

# 유사 연구(1)

## 사용자 후기를 중심으로한 글로벌 화장품 브랜드의 소비자 만족도 분석

1.글로벌 화장품 상위 50개 브랜드를 대상으로 온라인 후기를 수집.

2.감성분석(Sentiment Analysis)기법을 활용하여 화장품 브랜드들의 고객 만족도를 평가.

3. LDA기법을 활용하여 브랜드별 만족도 평가에서 만족도(긍정과 부정 의견들)를 유발하거나 영향을 미치는 주요 결정요인들을 분석하고 브랜드별 개선 방향을 제시.

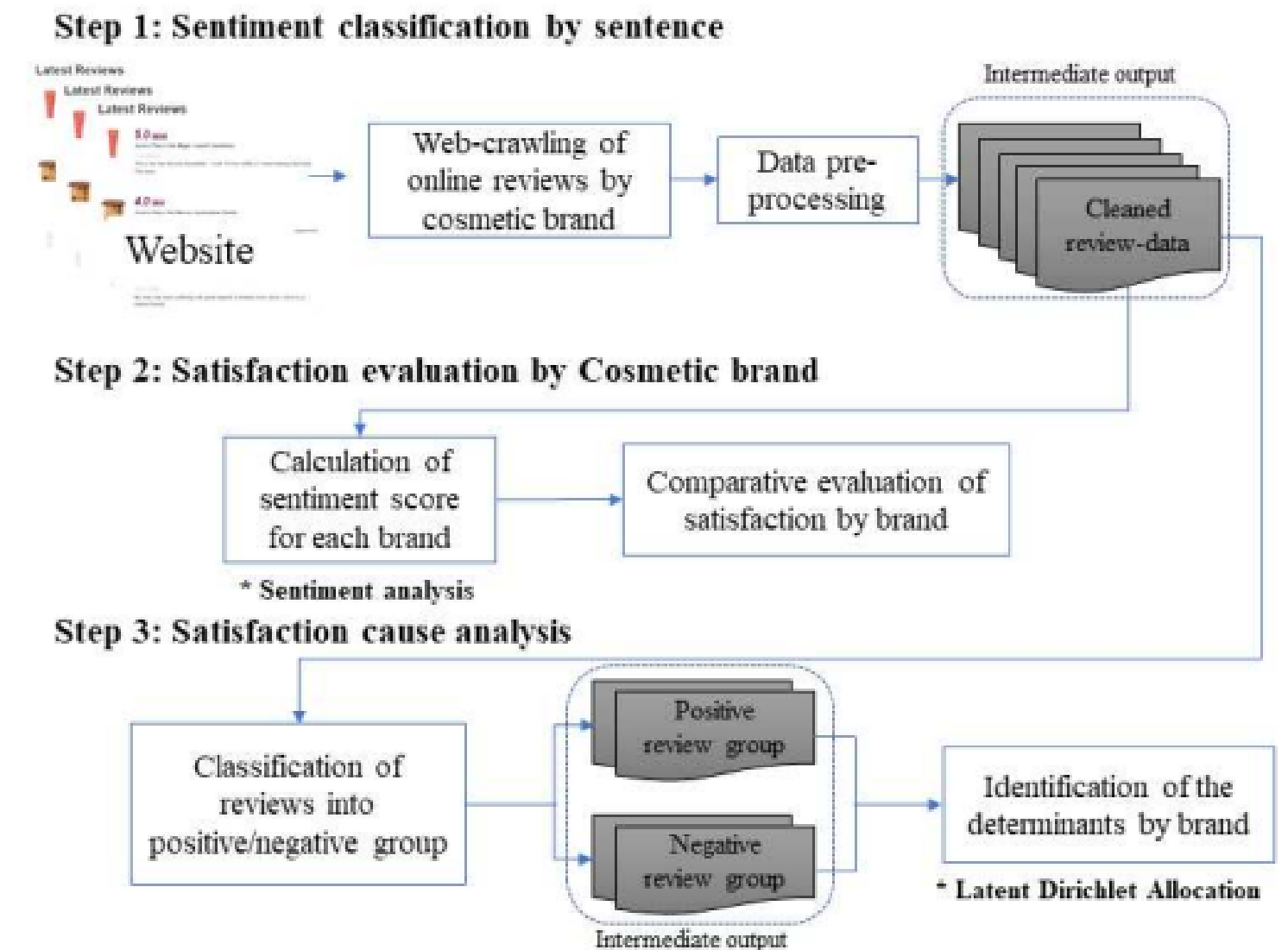


Figure 1. Online review-based satisfaction evaluation framework

박재훈(2021), 'Customer Satisfaction Analysis for Global Cosmetic Brands: Text-mining Based Online Review Analysis'.

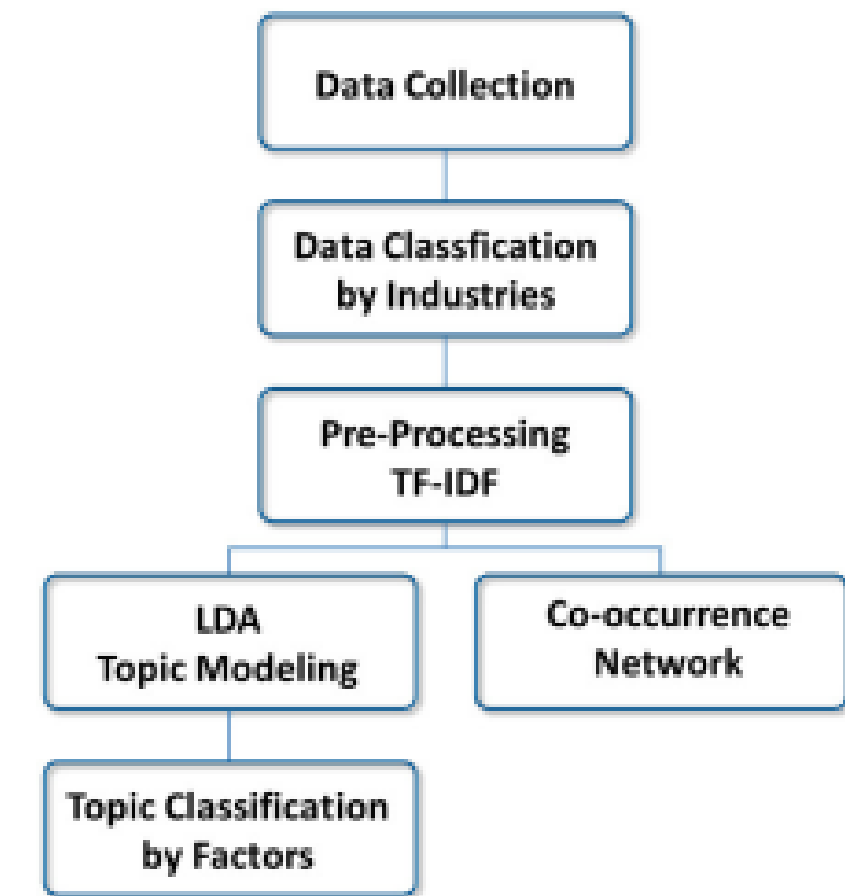
J Korean Soc Qual Manag 2021; Vol. 49, No.4: 595-607.

<https://www.jksqm.org/upload/pdf/jksqm-49-4-595.pdf>

# 유사 연구(2)

## 토픽모델링 기법을 활용한 산업별 직무만족요인 조사(잡플래닛 데이터 기반)

- 1.R을 이용한 웹 크롤링으로 기업 리뷰 수집.
- 2.R에서 제공하는 전처리와 'KoNLP'를 통해 명사, 형용사, 동사 형태로 추출.
- 3.LDA 토픽모델링 분석을 통해 토픽 추출 후, 30가지의 토픽으로 재분류.
- 4.출현 빈도가 높은 단어에 대해 동시 출현 분석을 통해 단어들 사이의 관계에 대해 연구 후 두 데이터 비교 분석.



〈Figure 2〉 Overall Research Process

# 유사 연구(3)

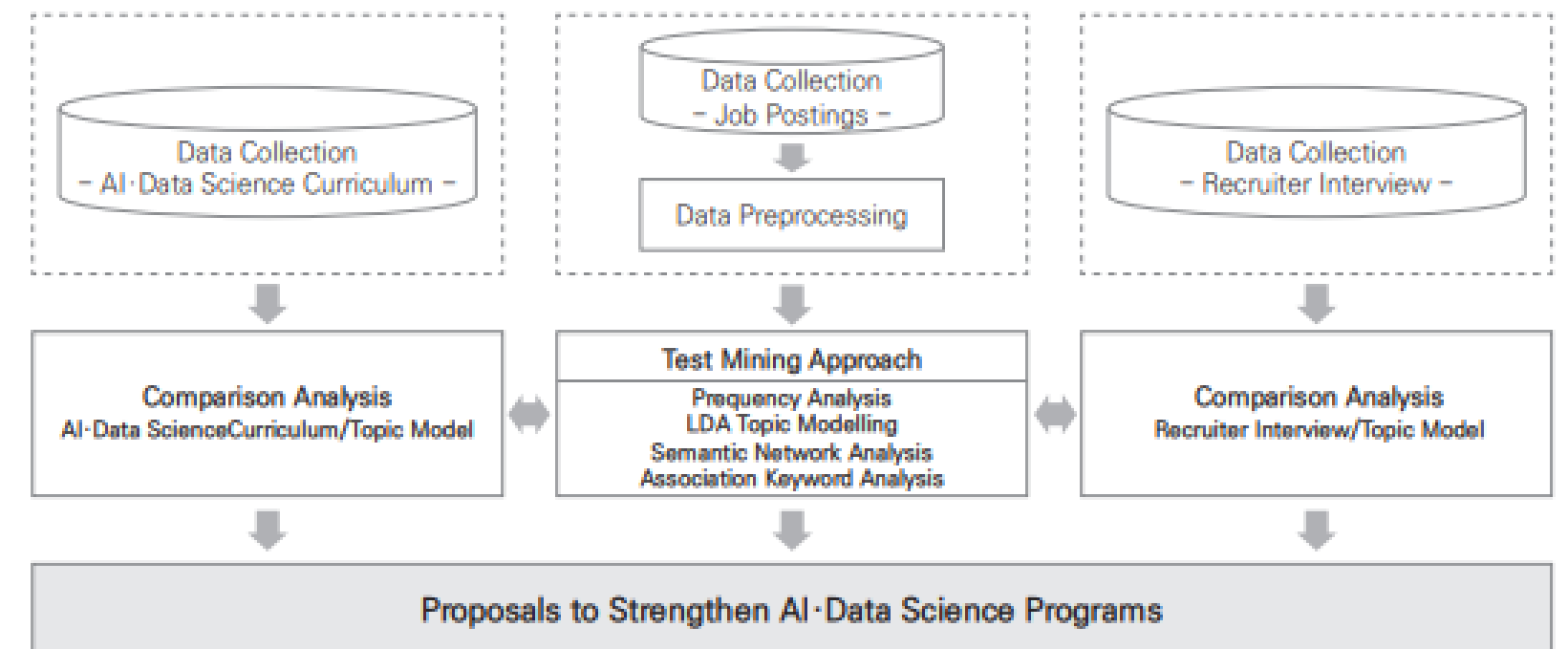
## AI·데이터 사이언스 분야 직무 역량 강화를 위한 커리큘럼 연구(LinkedIn)

1.Linkedin 사이트에서 텍스트 데이터 수집하고 전처리(토큰화, 불용어 제거, 명사 및 형용사 추출 등).

2. 전처리 완료된 데이터를 키워드 빈도 분석, LDA 토픽모델링, 연관어 분석

3.US NEWs Campuss에서 상위 5개 Data Science 분야 교과목 데이터 수집후 두 데이터에서 도출된 필요 역량 비교 분석.

4. 필요 역량 객관성 확인을 위해 Data Science분야에 재직중인 19인을 인터뷰.





감사합니다

