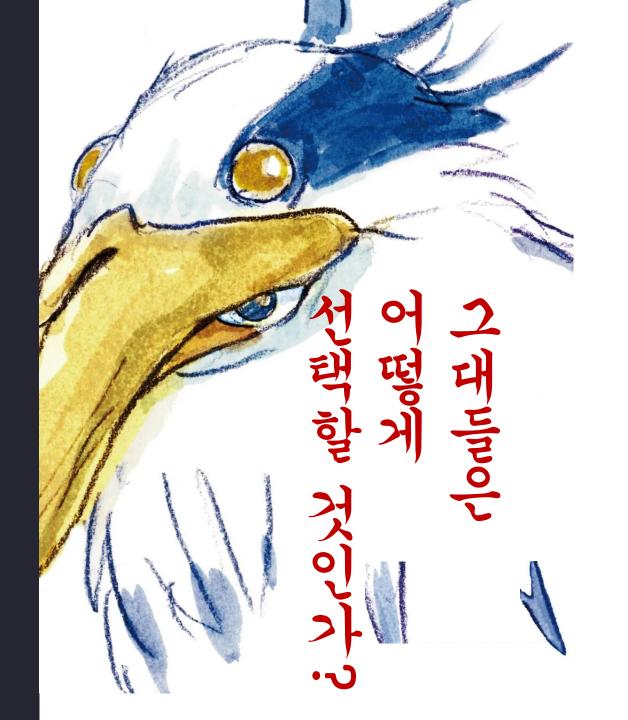


2021120118 김주연 2022111733 김하늘 2023111795 송진아 2022110821 조은빈



Contents

01

프로젝트 주제 소개

주제 선정 배경 주제 소개 및 문제 정의 02

데이터 수집

수집 기준 및 방법 수집 데이터 소개

03

데이터 전처리

전처리 과정 소개

04

데이터 시각화

기업별 사내문화 워드클라우드

05

결론 및 향후 연구 과제

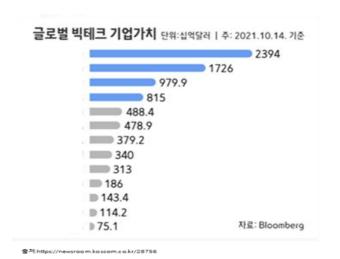
결론 및 한계점 극복방안 향후 연구 과제 소개

주제 선정 배경 소개

Why 조직문화?

조직 부합도 ↑ 조직 만족도 ↑ 스트레스 ↓

Why 빅테크 기업?





Why 텍스트 데이터?

조직문화 관련 다양한 내용이 담긴 텍스트 데이터

주제 소개 및 문제 정의

빅테크 기업 조직문화 분석



데이터 수집 기준 및 방법

취업 정보 사이트

"Connect Everything" # 카카오크루 에무도 가보지 않은 길을 함께 가는 사람들, 카카오크루는 세상의 문제를 새로운 시각과 방식으로 해결하고 싶고, 나이제 충분한 권한과 책임을 받아서 신나계 일하고 싶은 기업가정신이 중만한 사람들을 말합니다. # 일아는 방식 - 자기주도성, 명확한 권한과 책임을 가지고 일에 돌입합니다. - 광개 공유 - 빠른 실행을 위해 정보 및 의사결정 과정을 공개 공유합니다. - 국명 개유나계에선 : 최선의 의사결정을 위해 무엇이든 슬직하게 의견을 냅니다. **기업비전** **기업비전*** **기업보지 본문 연결을 통해 더 편리하고 즐거운 세상을 꿈뭡니다. 사업과 사원, 사업과 기술을 만증 가갑게 연결함으로써 세상을 어제보다 더 나은 곳으로 만들기 위해 노력하고 있습니다.

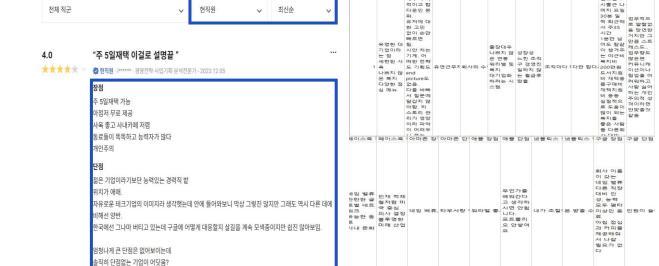
[현직자평가]

현진자 한마디 키워드

- 네이버: #복지 #자율성 #뛰어난동료 #보람 #성과중심 #과도한업무



- 수집 기준: 기업소개, 인재상, 사내문화 해시태그
- 수집 방법: 웹 크롤링
- 수집량: 30개(10*3)



블라인드

- 수집 기준: 현직원 필터링, 최신순 필터링, 장점리뷰, 개선점리뷰
- 수집 방법: 웹크롤링

주위가 다 능력자라 한없이 작아지는데 원동력으로 삼고 성장가능

- 수집량: 1000개(10개 기업 장점 50개, 단점 50개)

데이터 전처리 과정

1단계: 텍스트 교정

Row data 속 특수문자, 비속어 처리 (정규표현식 사용)

2단계: replace 딕셔너리

```
[네, 임밸류, 빠, 탄탄한, 글로벌, 네트워크, 빠, 유능한, 동료, 빠, ...
인터뷰, 없이, 팀, 변경, 빠, 내부, 오...
!, 보장, 빠, 좋은, 동료, 들]
커리어, 성장, 양, 날, 의, 검, 본인,...
아 플
좋은 워 밸 좋은 급여 좋은 임밸류
replace = {
'워라벨': '워라밸',
'워밸': '워라밸',
'워': '워라밸',
'라': '워라벨'.
```

Replace 딕셔너리로 합성명사 지정어 지정

3단계: 품사 태깅 및 토큰화

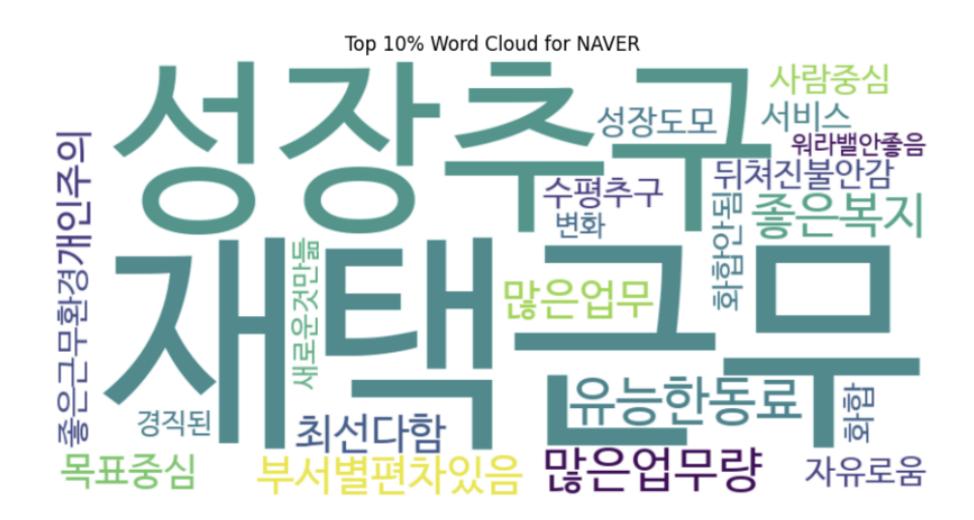
단어 토큰화: ['유명한', '대기업이라는', '점', '세련된', 품사 태깅: [('유명한', 'JJ'), ('대기업이라는', 'NNP'),



네이버 주식회사 검색 포털 국내 최초 어린이 유명 대기업 세련된 사옥 나쁘지않은

Customized KoNLPy로 형태소 분석 후 명사와 형용사만 토큰화 및 최종 정리

유지어터 그린이 N





Top 10% Word Cloud for KAKAO



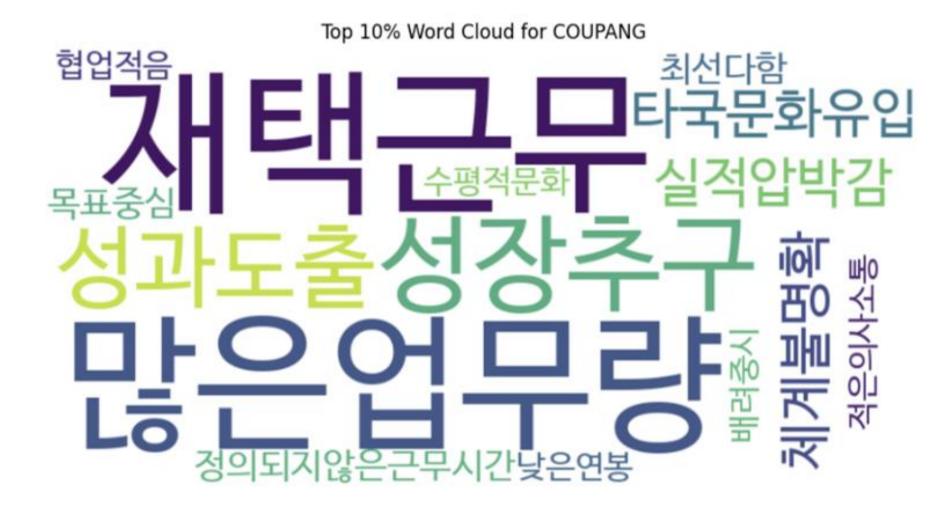
이불 속 완두콩 📭



Top 10% Word Cloud for LINE



불도저 태양이



따뜻한 아이스 아메리카노 배달이



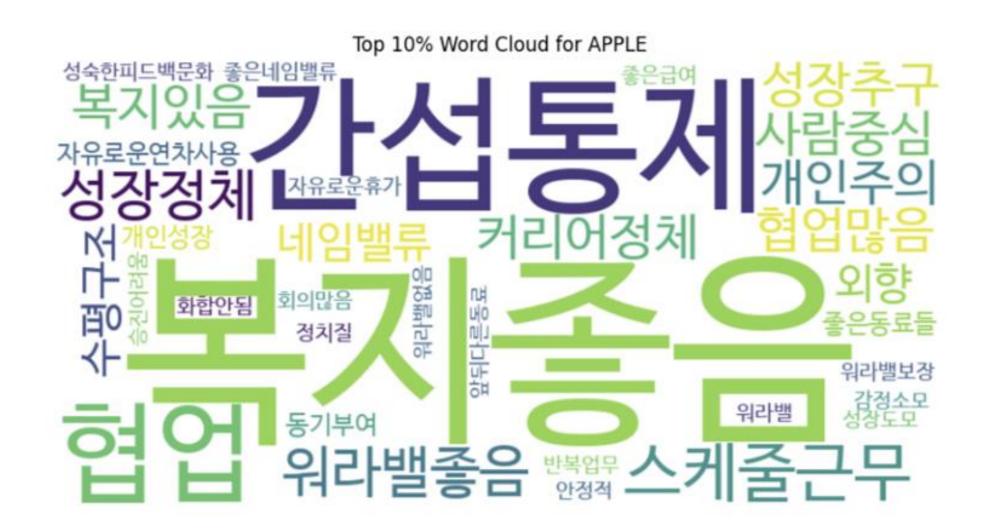
겉바속촉 파랑이 🚹



다함께 쑥쑥쑥 성장하는 정글이 @



베스킨라빈스 31 아삭이 🛊



만능 맥가이버 두둥



워커홀릭 G박사 **⑤**



결론 및 한계점 극복방안

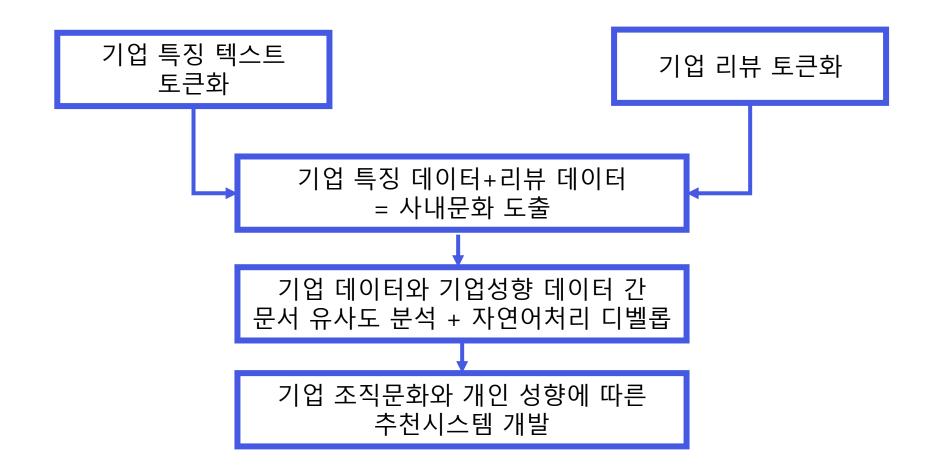
결론

- 기업별 조직문화 파악 가능
- 같은 산업군의 조직문화는 대체적으로 비슷하다는 것 파악
- 사내문화와 본인의 성향 비교하여 취업준비 판단 기준으로 활용 가능
- 형성한 데이터셋 기반으로 새로운 기업데이터 입력 시, 해당 기업 조직문화 파악 가능

한계점 극복방안

- 한계점: 익명 기업리뷰 데이터의 신빙성 이슈
- 극복방안
 - 1) 블라인드보다 신빙성을 보장할 수 있는 대체 플랫폼 서칭
 - 2) 익명이 아닌 철저한 인증을 거쳐 등록된 리뷰데이터로 대체

향후 연구 과제 소개



참고문헌

- 백혜연, "기업 리뷰 웹 사이트 텍스트 분석을 통한 직원 불만 표현 추출과 불만 원인 도출 및 해소 방안", 한국정보통신학회논문지, 2019.04
- 머스트 뉴스, "직장인, 이직·퇴사 주요 요인은 조직문화", 2023.01.29, http://mustnews.co.kr/View.aspx?No=2697804
- 전영근 외 4,,"기업 리뷰 텍스트마이닝을 통한 기업 분석 시스템 개발", 대한산업공학회, 2017.11
- 안병대 외 1, "기업성격이 직무만족에 미치는 영향: MBTI 조직성격유형과 잡플래닛 기업리뷰를 활용하여", 한국지능정보시스템학회, 2021.12