
비정형 데이터 2차 발표

재직자 리뷰 기반 고 선호도 기업의 만족요인 분석

Seoul Tech
Data science
2020.11.10
김인조
이경찬



서론

연구배경

- 최근 청년 실업난이 심화되면서 사회 문제로 대두되고 있지만 더불어 중소기업 및 스타트업 또한 인력난에 시달리고 있음
- 이는 소수의 복지가 좋은 대기업에만 구직자가 몰리기 때문으로 파악됨
- 청년들이 선호하는 기업의 리뷰를 분석하여 해당 기업을 선호하는 이유를 파악하면 인력난 해소에 기여할 수 있을 것으로 파악됨

연구동기

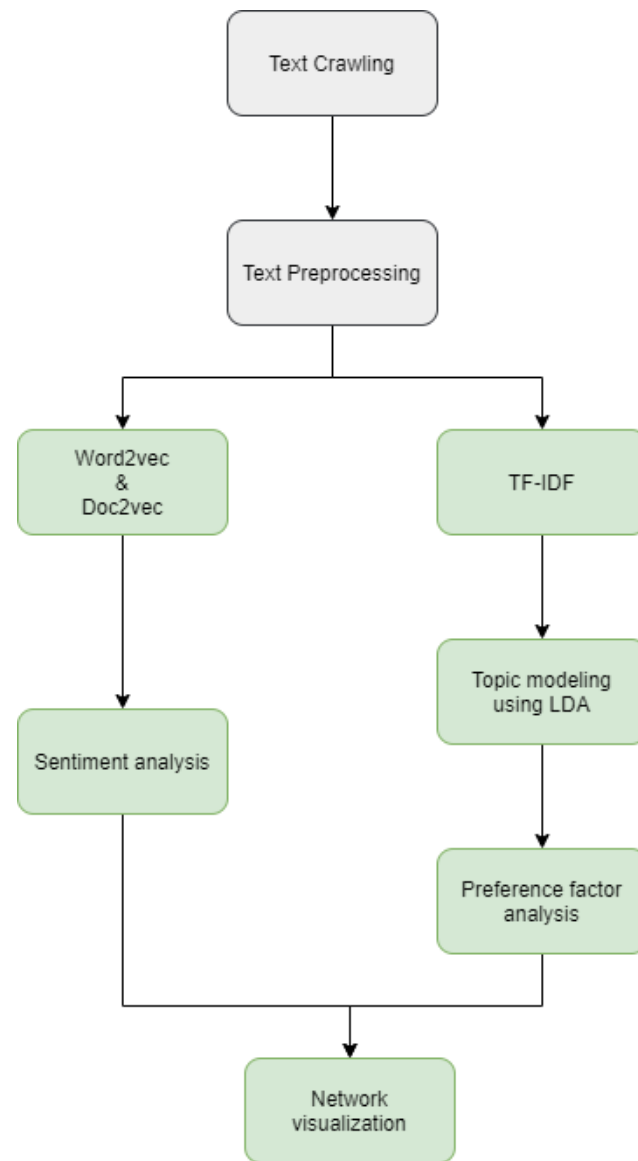
- 청년들의 기업 선호 요인에 대해 정량적으로 분석한 연구가 많지 않음
- 직원의 입장에서 진솔하게 기업을 평가한 리뷰 데이터를 바탕으로 해당 기업을 선호/비선호 하는 이유를 분석
- 기업 평가가 높은 기업의 만족요인과 불만족 요인을 대조군과 비교분석함으로써 나타나는 차이를 기업 입장에서 활용할 수 있음

연구 목적

- 기업 정보 리뷰 사이트 '잡플래닛' 에서 선호도가 높은 50개의 기업과 평점 2점대 기업 중 50개 기업의 리뷰 데이터를 비교하여 선호 요인을 분석하고자 함



연구 프레임 워크





연구 프레임 워크

감성 분석

- 감정이 명확한 단어들을 미리 labeling(좋다, 거슬린다)
- 단어들을 임베딩 후 해당 공간에서 거리를 기반으로 각 단어에 점수 부과

요인 분석

- 추출한 명사를 바탕으로 DTM(Document -Term Matrix) 생성
- LDA를 이용한 Topic modeling
- 추출된 토픽을 바탕으로 선호 요인 파악





진행 현황

데이터 수집 현황

분석 데이터

- 잡플래닛 내 IT 산업군 중 만족도 Top 50 기업 (만족도 3.6~5.0)
- 대조군으로 만족도 2점대인 기업 중 랜덤하게 50개의 기업을 샘플링
- 만족도 Top 50인 기업들을 하나의 실험군으로 2점대 기업을 대조군으로 설정

데이터 사전 처리

- 기업당 약 70-150개의 리뷰 데이터가 존재
- 삼성, SK와 같은 대기업은 700개 이상의 리뷰가 존재하여 데이터의 imbalance을 야기하고 선호 이유도 단순히 '대기업이기때문' 이므로 분석 대상에서 제외





진행 현황

데이터 수집 현황

기업명(평점)

틱톡(3.9)
라이엇 게임즈(4.4)
네이버 웹툰(4.3)
...

실험군(평점 Top50)

기업명(평점)

삼미정보시스템
(2.0)
와이즈스톤(2.0)
웹솔루스(2.0)
...

대조군(평점 2점대 50개 기업)

진행 현황

데이터 수집 현황

[기업리스트 페이지]

Jobplanet 채용 기업 뉴스 연봉 멤버십 79% 할인 중

Q 기업, 채용공고를 검색해보세요.

IT/웹/통신 2차 산업군 지역

IT/웹/통신 산업군의 1125 검색결과 총 만족도 순

- 1 페이스북 (유) IT/웹/통신 서울 54 기업리뷰 113 연봉정보 49 면접정보 총 만족도 순 ★★★★★ 4.6 평균 14,001 만원
- 2 (주)셀메이트 IT/웹/통신 서울 24 기업리뷰 15 연봉정보 3 면접정보 총 만족도 순 ★★★★★ 4.6 평균 3,075 만원
- 3 버즈빌(주) IT/웹/통신 서울 39 기업리뷰 24 연봉정보 63 면접정보 총 만족도 순 ★★★★★ 4.4 평균 3,751 만원
- 4 라이엇게임즈코리아 IT/웹/통신 해외 43 기업리뷰 37 연봉정보 43 면접정보 총 만족도 순 ★★★★★ 4.4 평균 6,337 만원

[개별 리뷰 Box]

총 54개 중 24개의 기업리뷰 IT/인터넷 전체 재직상태

③ IT/인터넷 ④ 전직원 ⑤ 서울 2020. 07. 31

② ★★★★★ ⑥ "개발 대우도 잘해주고 여러모로 근무하기 편한 분위기"

⑪ 승진 기회 및 가능성

⑦ 장점 개방적인 분위기에서 자유롭게 일할 수 있는 환경. 눈치보지 않을 수 있는게 제일 좋음

⑧ 단점 핵심 업무 외에 해결해야할 다른 업무들이 조금 있어서, 일에 온전하게 집중하기에는 어렵다

⑨ 경영진에 바라는 점 직원들의 요구 사항을 잘 들어주셨으면 좋겠습니다.

이 기업은 1년 후 성장하고 있을 것이다.

이 기업을 추천 합니다!

⑩ 🍎 도움이 돼요 0 📄 페이스북에 공유 신고하기

	①기업명	②총평점	③직종	④재직상태	⑤작성일	⑥리뷰title	⑦장점	⑧단점	⑨경영진에 바라는점	⑩도움	Score1	Score2	Score3	Score4	Score5
1	페이스북 코리아	5	IT/인터넷	전직원	2020. 07. 31	개발 대우도 잘해주고 여러모로 근무하기 편한 분위기	개방적인 분위기에서 자유롭게 일할 수 있는 환경. 눈치보지 않을 수 있는게 제일 좋음	핵심 업무 외에 해결해야할 다른 업무들이 조금 있어서, 일에 온전하게 집중하기에는 어렵다	직원들의 요구 사항을 잘 들어주셨으면 좋겠습니다.	0	5	5	5	5	5



진행 현황

Text preprocessing

POS Tagging

- 텍스트에서 명사만 남기고 모두 제거

Stopword

- <https://bab2min.tistory.com/544> 를 참고하여 stopwords 진행
- 한국어 단어의 평균 길이는 2-3정도로 추정되므로 길이가 1인 텍스트는 의미가 없다고 판단하고 제거

Text
핵심 업무 외에 해결해야할
다른 업무들이 조금 있어서,
일에 온전하게 집중하기 힘들다



Index	word	pos
0	핵심	Noun
1	업무	Noun
4	해결	Noun
7	다른	Noun
8	업무	Noun
11	조금	Noun
16	온전	Noun
18	집중	Noun