

원하는 기업, 다(多) 잡(job)아봐!

: MZ세대를 위한 기업 맞춤형 추천 서비스

TEAM 달콤챔피언

이수정 이은서 이정우 위재성 조혜현

목차

a table of contents

MZ세대 취업 현황 및 니즈 파악

서론

본론

결론

1. 활용 방안
 2. 추가 아이디어 및 서비스
 3. 최종 결론 (의의 및 한계)
1. 서비스 소개 및 목적
 2. 하드파워/소프트파워 정의
 3. 하드파워 – 트렌드지표 1 (재무제표)
 4. 소프트파워 – 트렌드지표 2 (기업 가치평가)
 5. 소프트파워 – 트렌드지표 3 (기업 인프라)

Part 1

서론

MZ세대 취업 현황 및 세태 파악

[MZ세대 열풍 리포트②]

90년대생 기업 들어왔다 떠난다...

'체질개선 필요'

60% 이상 입사 후 2년 안에 퇴사한다...

짧게 근무 후 퇴사하는 인원 급증

기업이 비전을 제시하는 쪽으로 가야 된다고 전문가 조언

'첫 이직 경험'

있다

75.5%

없다

24.5%

투데이코리아=안현준 기자 | 최근 기업에 90년대생들이 신입사원으로 입사하고, 80년 중·후반대생 임원이 등장하면서 대다수의 기업들의 문화가 급격하게 변화하고 있지만 짧게 근무 후 퇴사하는 인원들이 더 많아 실질적인 기업들의 체질 변화가 필요하다는 목소리가 나오고 있다.

[출처: 잡코리아]

특히 극심한 구인난과 취업난이 공존하는 현상까지 발생하면서 근본적인 해결이 필요하다는 목소리도 거세지고 있다.

통계청이 발표한 11월 고용동향에 따르면 지난달 15세 이상 취업자는 1년 전보다 55만3000명 증가했지만, 30대 취업자 수는 6만 9000명 감소해 21개월 연속 감소세를 보인 것으로 나타났다.

또한 지난해 청년층 실업자 수는 37만명으로 집계되었으며, 추가 취업을 희망하는 시간 관련 추가 취업 가능자 수는 14만9000명으로 2016년 조사때보다 2배 증가한 것으로 집계됐다.

하지만 취업 후에도 상당수의 청년들은 1년도 지나지 않아 퇴사하는 것으로 알려졌다.

MZ세대 취업 현황 및 세태 파악

- 극심한 취업난을 겪고 있는 MZ세대.
그토록 원하던 취업에 성공하여 꽃길만 걸을 것 같은 그들은 1년도 지나지 않아 퇴사하는 것으로 밝혀짐.
- 즉 60% 이상이 입사 후 2년 안에 퇴사한다는 현황에서 기업의 인사 담당자들은 조직에 얼마나 오랫동안 머무를 수 있는지를 가장 중요하게 보게 되었다며 힘든 점을 토로함. 하지만 청년들은 기업부터 바뀌어야 한다고 주장했으며 회사 내의 자세한 사정과 예상과 동떨어진 회사 생활에 회의감을 느껴 퇴사했다고 고백함.
- 또한, MZ세대의 근간을 이루는 90년대생들은 연봉이라는 대외적인 보상보다는 비전, 개인의 성취와 성과를 중요시 하기에 이를 반영할 수 있는 기업 트렌드 지표가 더욱 절실한 셈임.

MZ세대 퇴사 증가

극심한 구인난 + 취업난

기업 트렌드 지표 절실

Part 2

본론

서비스 소개 및 목적

서비스 소개

MZ세대들의 회사 정착율을 높이기 위해 필요한 해결방안은 무엇인가?
그 답은 기업에 대해 흔히 얻을 수 있는 정보인 연봉, 복지 이외의 트렌드 지표들을 반영해
맞춤형 기업을 추천해주는 서비스가 될 것임. 이에 구직자를 위한 새로운 기업 트렌드 지표를 발굴해
구직자의 성향과 니즈를 파악할 수 있는 서비스를 제안하고자 함.

서비스 목적

MZ세대들이 구직을 할 때에, 본인의 니즈와 가치 및 비전에 걸맞는 기업을 최초로 추천 받아
퇴사를 하거나 또다시 힘들게 이직을 준비해야하는 행태를 감소시키는 데에 있음.
구직을 할 때 한번에 걸맞는 기업들을 골라내면 퇴사율을 줄여
기업과 구직자 모두에게 도움이 될 수 있음.

원하는 기업, 다(多) 잡(job)아봐!
: MZ세대를 위한 기업 맞춤형 추천 서비스

하드파워/소프트파워 정의

하드파워

"외적인 보상"

보상제도

복지제도

기업 안정성

채용 합리성

소프트파워

"내재적 동기"

개인 가치관과의
적합성

기업 이미지

기업 트렌드 지표

하드파워

1. 재무제표

안정성

수익성

성장성

활동성

시장가치비율

선호 반영하여 관심 업종
기업 **Top10** 추천

소프트파워

2. 기업 가치평가

기업 분위기 + 추구하는 가치

선호 기업 주요 이슈 키워드

선호 업종 주요 이슈 키워드

소프트파워

3. 기업 인프라

기업 근처 대중교통 이용 가능성

기업 근처 대중교통 이용 편의성

접근성 지수 도출

하드파워-트렌드지표 1 (재무제표)

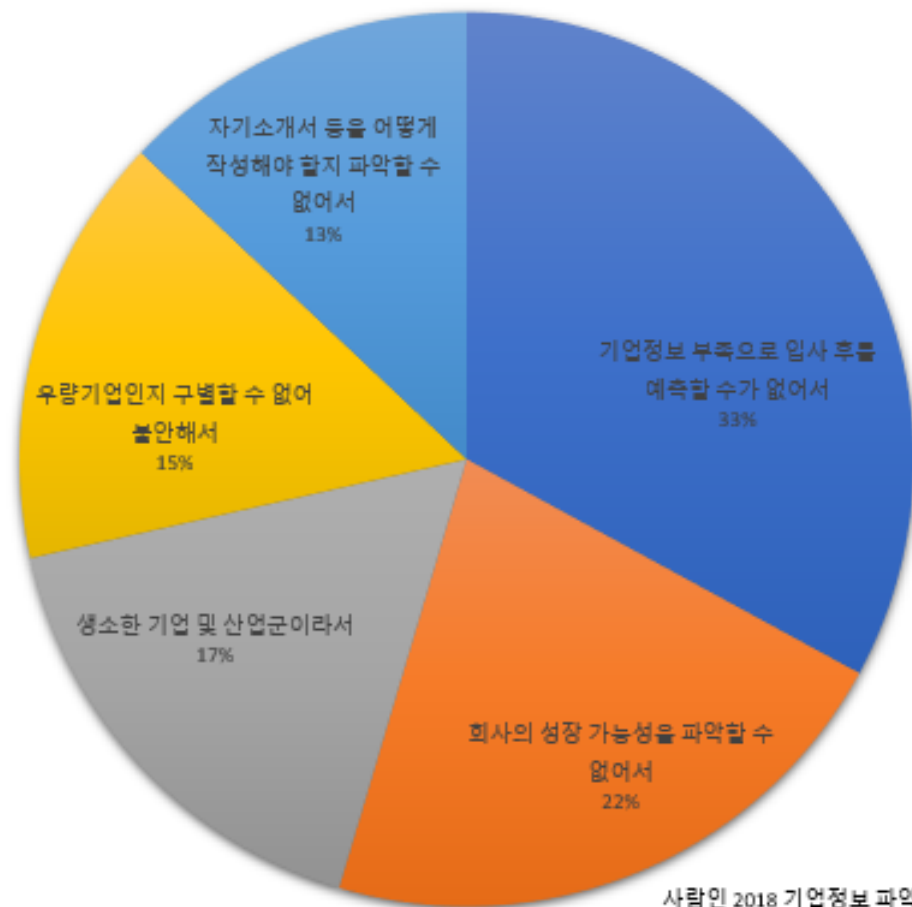
구직자를 위한 기업 재무제표 정보제공의 필요성

구인구직 매칭 플랫폼 사람이 구직자들을 대상으로 입사지원 포기사유를 조사한 결과, “기업정보 부족으로 입사 후를 예측할 수 없어서”가 33%, “회사의 성장 가능성을 파악할 수 없어서”가 22%, “우량기업인지 구별할 수 없어서”가 15%로, 입사지원 포기의 주된 이유가 됨.

기업의 다양한 정보와 성장가능성, 우량기업 등은 재무제표 분석을 통해 파악할 수 있으며, 이는 구직자 대상으로 입사지원시 기업의 재무제표 정보를 제공하는 것이 중요함을 시사함.

따라서 기업의 재무특성을 전공자가 아닌 일반인도 이해하기 쉽도록 분석 및 시각화하여 구직자들에게 제공할 필요가 있음.

입사지원 포기 사유



사람인 2018 기업정보 파악도 조사결과

[출처: 사람인]

하드파워-트렌드지표 1 (재무지표)

데이터 수집



>>

- ① 재무상태표
- ② 손익계산서
- ③ 포괄손익계산서
- ④ 현금흐름표
- ⑤ 주가 및 상장주식수

데이터 전처리

1. 동의어인 재무항목들을 하나의 이름으로 통일
2. 기업 개수가 1개 또는 2개인 업종 제거

>>

재무지표 선정 및 계산

기업의 대표적인 5가지 재무특성에 대한 재무지표 선정 및 계산

시장가치비율 재무지표

1. PER
2. PBR

안정성 재무지표

1. 부채비율
2. 유동비율
3. 고정비율

수익성 재무지표

1. 매출액영업이익률
2. 매출액세전순이익률

성장성 재무지표

1. 매출액증가율
2. 영업이익증가율

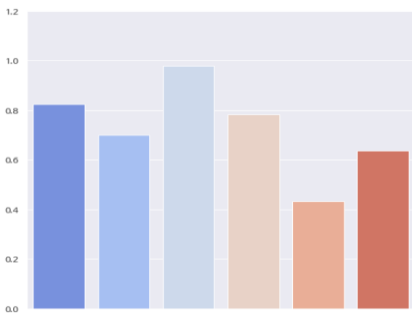
활동성 재무지표

1. 총자산회전율
2. 고정자산회전율

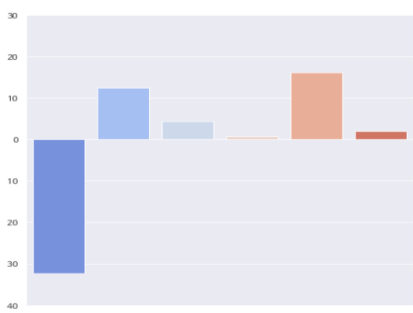
하드파워-트렌드지표 1 (재무제표)

재무특성지수 산출

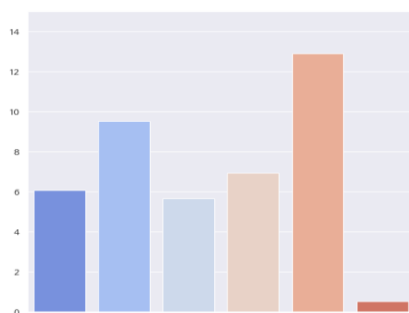
1. 업종별 표준화



업종별 총자산회전율 평균



업종별 PER 평균



업종별 매출액영업이익률 평균

>> 업종별로 재무지표 평균이 다르므로
기업별 모든 재무지표를 업종별로 표준화

ex)

$$A\text{기업의 표준화된 } k\text{번째 안정성 재무지표 } I_{A,k} = \frac{(S_{A,k} - \text{mean}_{A,k})}{\text{std}_{A,k}}$$

$S_{A,k}$: A기업의 k번째 안정성 재무지표,
 $\text{mean}_{A,k}$: A기업이 속한 업종의 k번째 안정성 재무지표 평균
 $\text{std}_{A,k}$: A기업이 속한 업종의 k번째 안정성 재무지표 표준편차

2. 가중치 선정

업종별로 중요하게 보아야 할 재무지표가 다르므로 안정적인 재무특성지수 산출을 위해 같은 재무특성을 나타내는 재무지표는 가중치의 절댓값을 동일하게 설정.
부채비율, 고정비율은 낮을 수록 좋은 재무지표이므로 가중치에 (-) 붙임.

안정성

1. 부채비율 가중치: -0.33
2. 유동비율 가중치: 0.33
3. 고정비율 가중치: -0.33

수익성

1. 매출액영업이익률 가중치: 0.5
2. 매출액세전순이익률 가중치: 0.5

성장성

1. 매출액증가율 가중치: 0.5
2. 영업이익증가율 가중치: 0.5

활동성

1. 총자산회전율 가중치: 0.5
2. 고정자산회전율 가중치: 0.5

시장가치비율

1. PER 가중치: 0.5
2. PBR 가중치: 0.5

>>

3. 재무특성지수 산출

재무에 대해 잘 알지 못하는 구직자에게도 효율적으로 정보를 제공하기 위해
5가지 재무특성을 나타내는 지수를 산출.

모든 재무지표를 업종별로 표준화한 뒤 앞서 설정한 가중치를 곱하고 같은 재무특성을 나타내는 값들을 모두 합하여 재무특성지수 산출.

ex)

$$A\text{기업의 안정성 지수} = \sum_{k=1}^n I_{A,k} * W_k$$

$I_{A,k}$: A기업의 표준화된 k번째 안정성 재무지표,
 W_k : k번째 안정성 재무지표의 가중치, n: 안정성 재무지표 개수

하드파워-트렌드지표 1 (재무제표)

AHP 분석법

행정학과 정책학의 분석기법 중 하나로,
다수 속성들의 상대적 중요도를 파악함으로써 최적의 대안을 선정하는 기법
일관성 검증을 통해 전이적 일관성이 얼마나 유지되고 있는지 검증

알고리즘

1. 5가지 재무특성 (안정성, 수익성, 성장성, 활동성, 시장가치비율)을 두 개씩 짝지어 개별 구직자가 생각하는 중요도를 1~9 사이 숫자로 입력 받고 쌍대비교행렬 구성

쌍대비교행렬

$$\begin{bmatrix} 1 & a & b & c & d \\ \frac{1}{a} & 1 & e & f & g \\ \frac{1}{b} & \frac{1}{e} & 1 & h & i \\ \frac{1}{c} & \frac{1}{f} & \frac{1}{h} & 1 & j \\ \frac{1}{d} & \frac{1}{g} & \frac{1}{i} & \frac{1}{j} & 1 \end{bmatrix}$$

(a, b, ..., j : 해당 구직자가 입력한 값)

CR ≥ 0.2

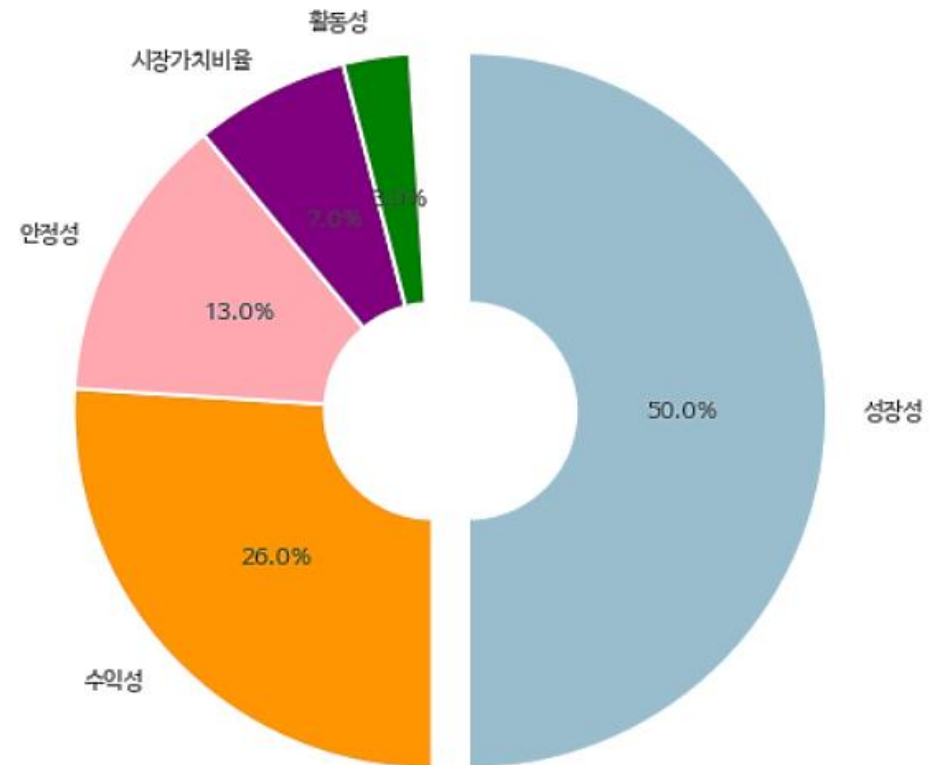
$$2. \text{일관성 비율 } CR = \frac{\lambda_{\max} - n}{(n-1) * \text{Random Index}}$$

(λ_{\max} : 쌍대비교행렬의 최대 고유값, n : 쌍대비교행렬의 열 또는 행의 개수)

CR < 0.2

3. 쌍대비교행렬의 각 요소들을 열별 합으로 나눈 뒤 행을 따라서 더함으로써 해당 구직자의 5가지 재무특성에 대한 상대적 선호도 산출

재무특성 선호도



하드파워-트렌드지표 1 (재무제표)

알고리즘

4. 행은 기업명, 열은 5가지 재무특성으로 하는 재무특성지수 행렬과 AHP 분석법을 통해 산출한 개별 구직자의 재무특성 선호도 벡터를 행렬곱하여 구직자 맞춤형 기업 선호도를 산출

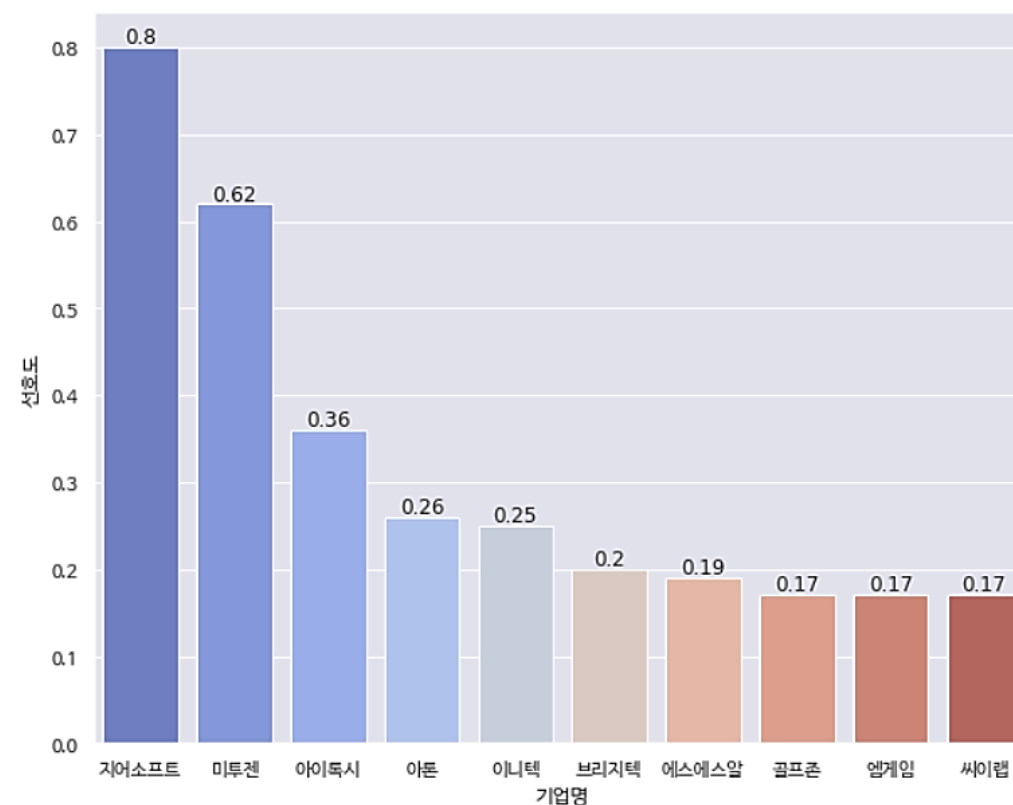
<해당 구직자의 기업별 선호도>

$$\begin{array}{c} \text{기업명} \end{array} \begin{array}{c} \text{재무특성} \\ \begin{bmatrix} 0.06 & \cdots & -0.04 \\ \vdots & \ddots & \vdots \\ 0.10 & \cdots & 0.01 \end{bmatrix} \end{array} \begin{array}{c} \begin{bmatrix} 0.5 \\ 0.26 \\ 0.13 \\ 0.07 \\ 0.03 \end{bmatrix} \end{array} = \begin{array}{c} \begin{bmatrix} 0.05 \\ -0.07 \\ \vdots \\ 0.07 \\ -0.31 \end{bmatrix} \end{array}$$

재무특성지수 행렬
구직자의 재무특성 선호도 벡터
구직자의 기업별 선호도

5. 기업 선호도 Top10 기업을 해당 구직자에 대한 추천기업으로 선정

기업 선호도 Top10



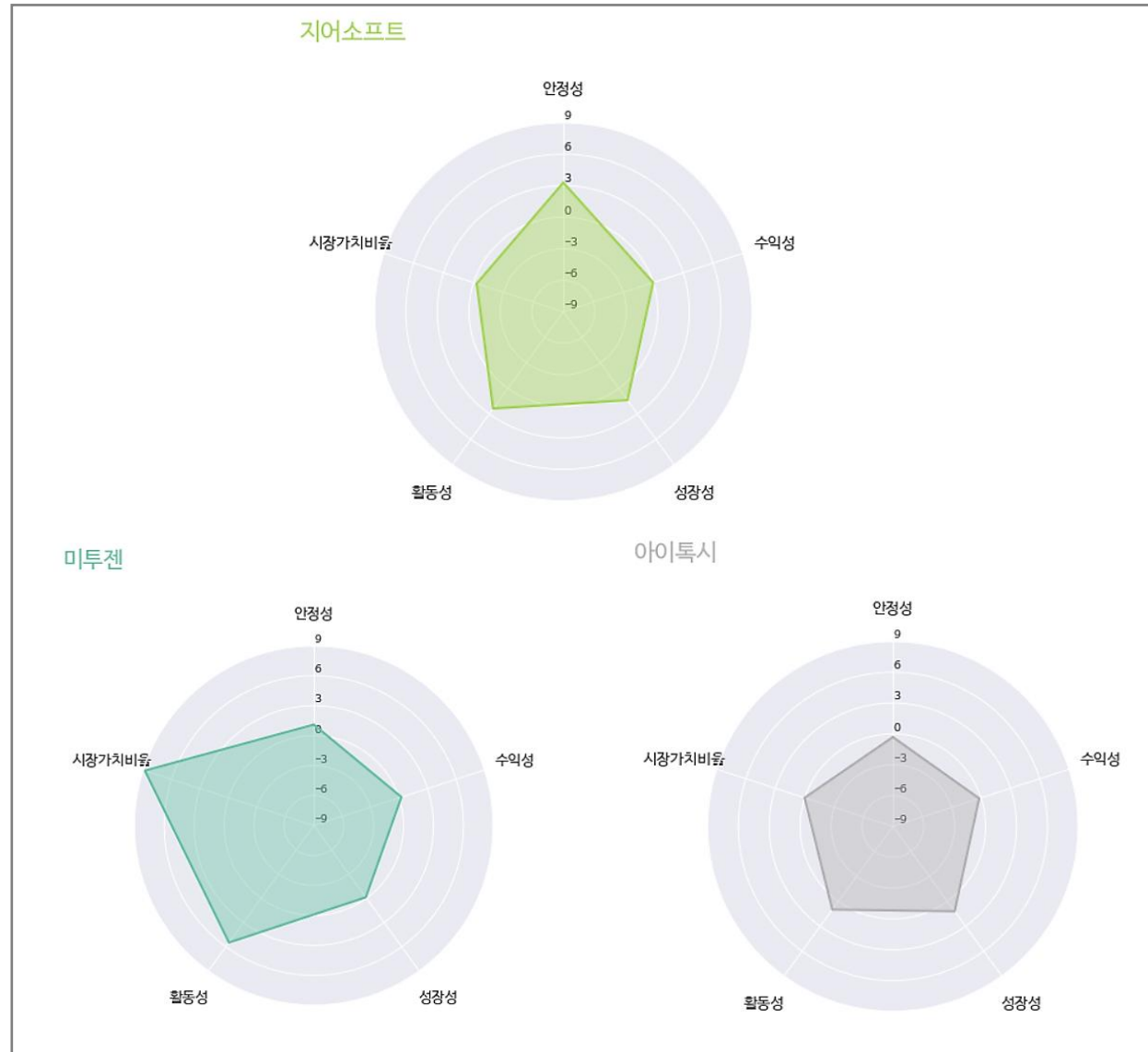
하드파워-트렌드지표 1 (재무제표)

기업선호도 Top3 기업의 재무특성 레이더차트

해당 구직자에게 추천된 기업들의 재무특성을 시각화하여, 추천된 기업의 특징에 대한 정보를 구직자에게 제공함.

재무 지식이 없는 구직자도 기업들이 어떠한 점에 특화되어 있는지를 파악할 수 있도록 하여 추천 기업에 대한 정보를 효율적으로 제공함.

가상의 구직자에게 추천된 세 기업 모두 5가지 재무특성이 균형 잡혀있으며, 그 중 지어소프트는 안정성과 활동성이 높은 모습을 보이고 미투젠은 시장가치비율과 활동성이 높은 모습을 보임.



소프트파워 - 기업 가치평가

MZ세대, 기업의 성장성&비전 부족하면 조기퇴사 ↑

기업의 가치를 제시하여
구직자가 적합한 기업을 선택하는 것이 중요함

'MZ 조기 퇴사 이유'

개인 만족 우선

60.2%

그외

39.8%

[출처: 사람인]

기업 분위기 키워드

기업 비전, 인재상 제시

기업 최신 이슈

기업의 성장, 관심사 제시

업종 최신 이슈

관련 산업 성장, 동향 제시

기업 가치평가 Work Flow

기업 및 업종 뉴스

토픽 모델링

Topic Coherence와
Perplexity 모두 고려하여
토픽의 개수를
그리드 서치로 튜닝



Sklearn 라이브러리의
LDA 모델로
토픽 모델링 진행

>>

기업 비전과 분위기

키워드 추출

TF-IDF
자연어 데이터로부터
Feature를 뽑아내는 방법

>>

워드 클라우드

기업 분위기
한눈에 파악할 수 있도록
워드 클라우드 생성

데이터 수집

- 기업 1년치 뉴스
(2021.1.1 ~ 2021.12.31)
- 업종 최신 뉴스
(2021.12.27 ~ 2021.12.31)
- 기업 홈페이지
(경영 철학, 비전,
인재상, 기업 소개)

>>

데이터 전처리

- ① 특수 기호 전처리
정규표현식 이용하여
기사 내 특수 기호 제거
- ② 토큰화
형태소 분석기 okt 이용
형태소 단위로 토큰화
- ③ 불용어 제거
기사 내에서 분석에
필요 없는 불용어들 제거

소프트파워 - 트렌드지표 2 (기업 가치평가)

1. 기업 분위기 키워드

[기업별 분위기 키워드 필요성]

- 홈페이지 기업소개, 인재상, 경영철학 등에서 키워드를 추출하여 해당 기업에 대한 기본적인 정보 제시
- 기업별 구직에 영향을 미치는 기업의 이미지, 관심사와 구직자의 가치관, 관심사를 비교하여 구직자가 적합한 기업을 선택하도록 도움



VS



[미투젠]

글로벌 카지노 게임 개발사
전세계에 적극적인 게임 출시를 하는
기업이기에 해외시장에 관심을 가지는
구직자에 적합

[아이톡시]

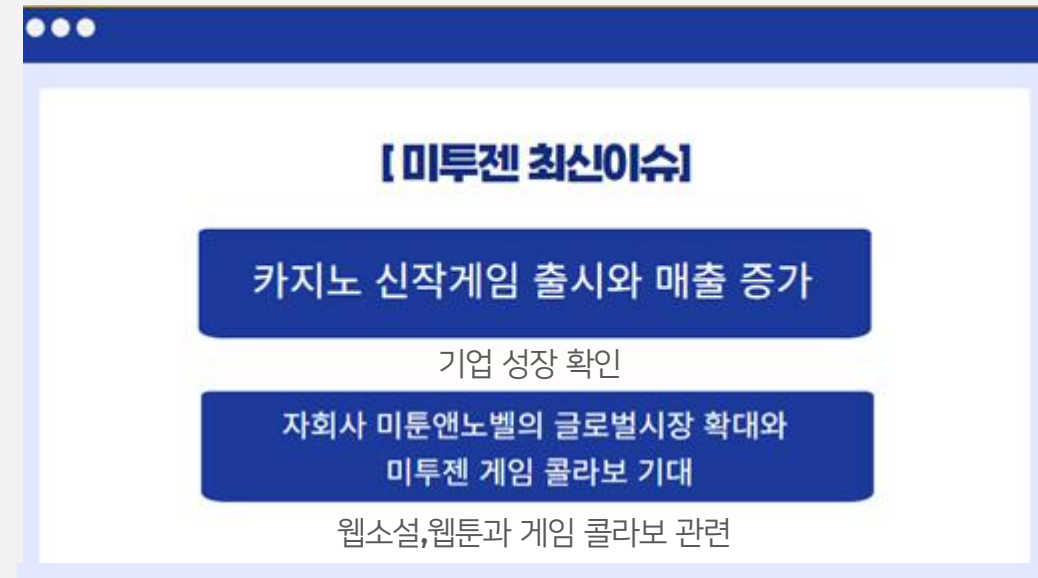
국내 IP 기반의 게임 개발 및 운영사로
최근에는 콘텐츠 사업으로 영역을 확장
다양한 콘텐츠 사업에 관심을 가지는
구직자에 적합

소프트파워 - 트렌드지표 2 (기업 가치평가)

2. 기업 최신 이슈 소개

[기업별 최신 이슈 필요성]

- 구직자가 관심있는 기업의 최근 1년의 뉴스 데이터로 토픽 모델링을 진행하여 해당 기업의 성장, 관심사 제시
- 해당기업에 대한 방대한 자료 중에서의 중요한 토픽, 이슈를 한 눈에 파악 가능
- 구직자는 관심기업의 최신 이슈를 통해 자소서나 면접 예상질문 준비 가능

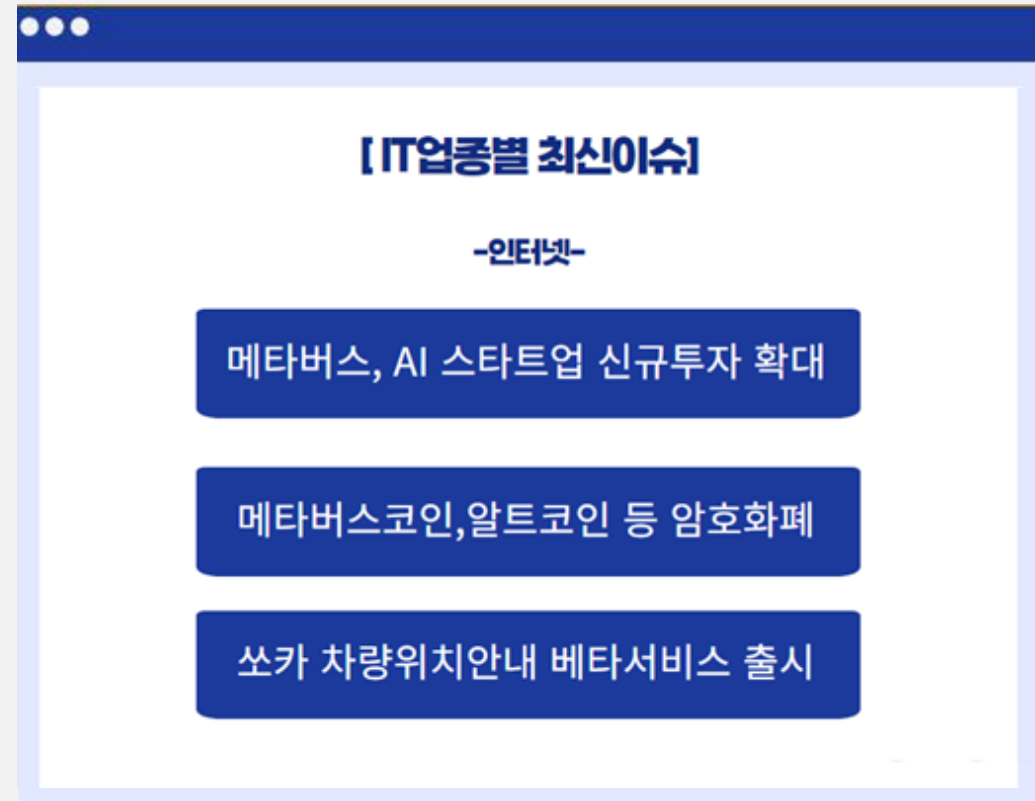


소프트파워 - 트렌드지표 2 (기업 가치평가)

3. 업종별 최신 이슈

[업종별 최신이슈 필요성]

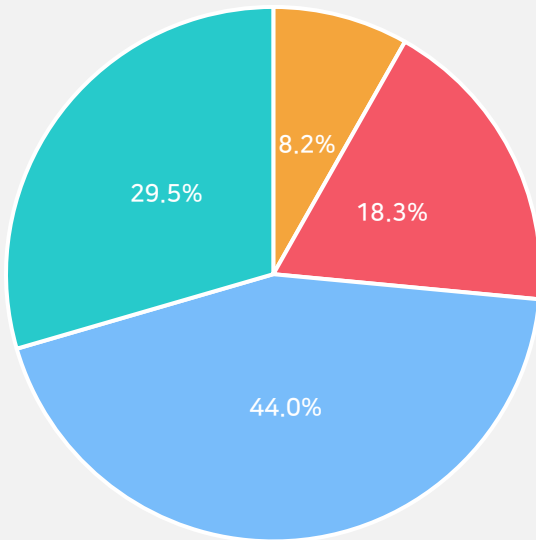
- 취업 준비를 위해 해당 기업의 최신 이슈 뿐만 아니라 해당 산업, 업종별 최신이슈 파악 필요
- 구직자가 관심있는 업종별 최신 일주일의 뉴스 데이터로 토픽 모델링을 진행하여 해당 업종의 최근 동향 제시
- 해당 업종에 대한 방대한 자료 중에서의 중요한 토픽과 이슈를 한 눈에 파악 가능



소프트파워 - 트렌드지표 3 (기업 인프라)

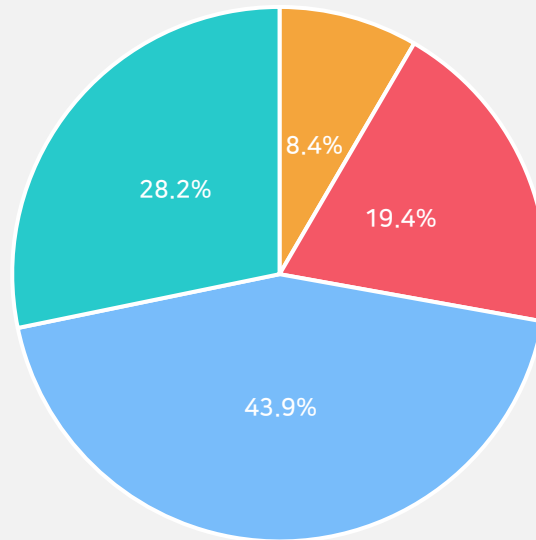
코로나 전후 출근 시간대 비교

2019년 출근 시간대 비율



■ 06-07시 ■ 07-08시 ■ 08-09시 ■ 09-10시

2021년 출근 시간대 비율



■ 06-07시 ■ 07-08시 ■ 08-09시 ■ 09-10시

코로나19 사태와 탄력근무제에도 불구하고
두 연도의 출근 시간대 비율 변화 크게 없음.

비대면 문화가 확산되고 출근 복지제도가 도입되었지만
이를 활용하는 기업이 그리 많지 않고 여전히 회사로
출퇴근하는 사람은 많다고 볼 수 있음.

이에 기업의 교통 접근성 지표가
여전히 기업을 고려함에 있어 중요한 지표라 할 수 있음.

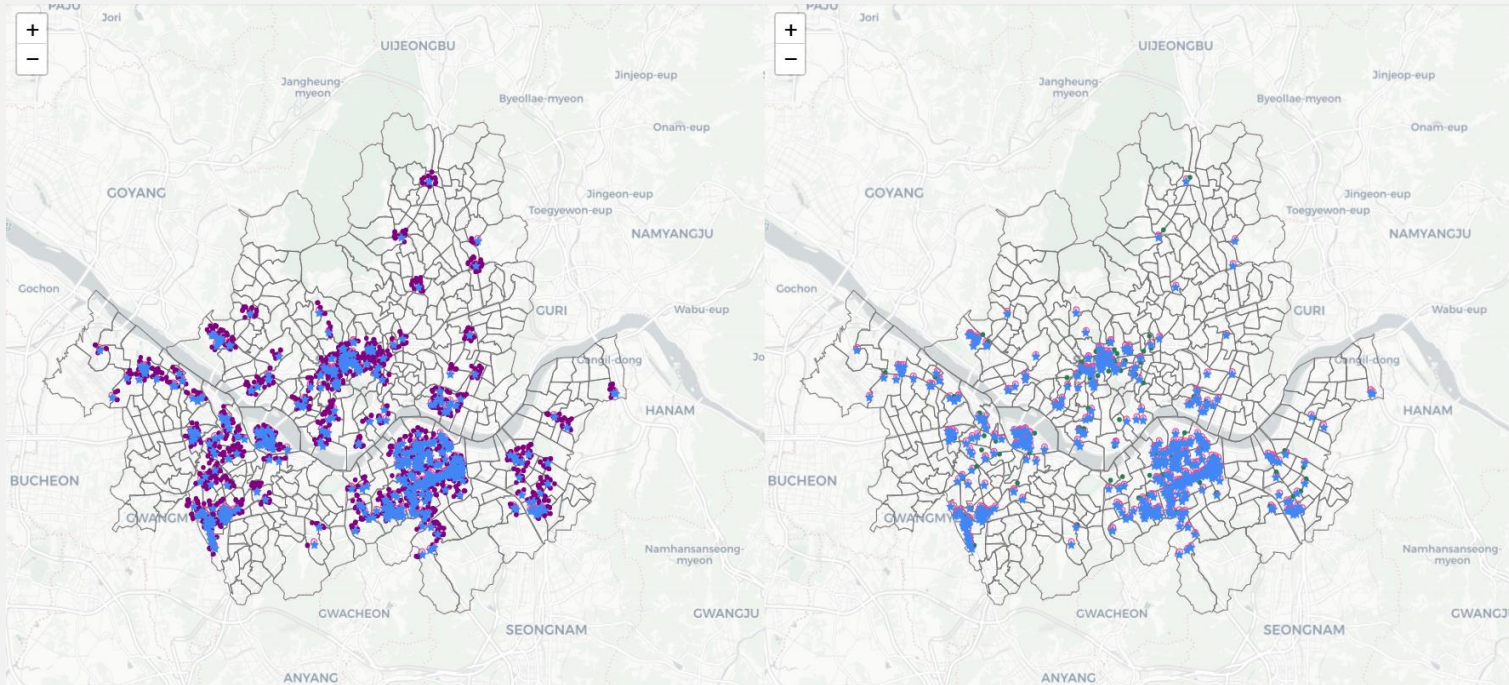
소프트파워 - 트렌드지표 3 (기업 인프라)

대중교통 접근성 평가 기준

미국의 대중교통 용량 및 서비스 편람(이하 TCQSM)에 따르면 대중교통 서비스 평가 기준을
이용 가능성과 이용 편의성으로 분류

버스정류장

지하철역



- TCQSM에 의하면 대중교통 시설의 영향권은 도보를 기준으로 버스 정류장은 8분 이내, 지하철역은 12분 이내로 가정
- 분석의 편의성을 위해 도보 10분 정도의 거리 500m로 영향권을 설정
- 기업별로 영향권 내의 대중교통 시설 데이터를 수집

기업 인프라 Work Flow

데이터 수집

대중교통 서비스 평가 기준

- 이용 가능성 {
- ① 버스 정류장, 지하철역 개수
 - ② 정류장 별 버스 대수, 역별 환승역 개수
 - ③ 정류장 및 지하철역까지의 거리
- 이용 편의성 {
- ④ 출퇴근 시간 승하차 인원
 - ⑤ 배차 간격
 - ⑥ 운행 대수 및 운행 거리

>>

상관분석

상관 Plot을 통한
최종 변수 선정

>>

클러스터링

모델링에 필요한
종속변수 값 산출

모델링

딥러닝 Sequential
모델 사용



SHAP value를 통한
가중치 계산



기업별
최종 접근성 지수 산출

소프트파워 - 트렌드지표 3 (기업 인프라)

데이터 수집

이용 가능성

시간적, 공간적으로 이용 가능한가를 나타내는 지표

- ① 버스 정류장, 지하철역 개수
- ② 정류장 별 버스 대수, 역별 환승역 개수
- ③ 정류장 및 지하철역까지의 거리

이용 편의성

얼마나 편리하게 이용할 수 있는가를 나타내는 지표

- ④ 출퇴근 시간 승하차 인원
- ⑤ 배차 간격
- ⑥ 운행 횟수 및 운행 거리

기업명	버스 변수							
	정류장 개수	버스 개수	접근거리	배차간격	운행거리	운행횟수	승차인원	하차인원
(주)세아홀딩스	17	15	287.0123	15.2741	41.9093	105.0645	1446	1510
(주)세아베스틸	17	15	287.0123	15.2741	41.9093	105.0645	1446	1510
(주)세아제강지주	17	15	287.0123	15.2741	41.9093	105.0645	1446	1510
애경케미칼(주)	11	48	289.9251	9.8640	43.2935	125.8155	2017	2042

기업명	지하철 변수							
	역사 개수	환승역 개수	접근거리	배차간격	운행거리	운행횟수	승차인원	하차인원
(주)세아홀딩스	1.0	2.0	165.3649	3.00	48.30	423.5	201968	219309
(주)세아베스틸	1.0	2.0	165.3649	3.00	48.30	423.5	201968	219309
(주)세아제강지주	1.0	2.0	165.3649	3.00	48.30	423.5	201968	219309
애경케미칼(주)	1.0	2.0	0	8.75	34.15	273.0	198407	222346

소프트파워 - 트렌드지표 3 (기업 인프라)

상관분석

상관관계가 높은 변수들끼리 결합해 최종 변수 선정

- 승차인원과 하차인원은 승하차인원으로 결합
- 승하차인원 = (승차인원 + 하차인원) / 2

단순한 정류장 개수나 역사 개수 보다는 노선의 다양성이

대중교통 접근성 판단에 더 용이하다 생각

- 정류장 개수와 버스 개수는 버스 개수를 대표 변수로 설정
- 역사 개수와 환승역 개수는 환승역 개수를 대표 변수로 설정

TCQSM에 따르면 일반적인 대중교통 서비스 평가에서

배차간격은 중요한 요소로 평가되고 있음.

- 배차간격을 제외한 운행거리, 운행횟수를 결합
 - 운행비율 = 운행횟수 / 운행거리

	버스							
	정류장개수	버스개수	*1/접근거리	*1/배차간격	운행거리	운행횟수	*1/승차인원	*1/하차인원
정류장개수	1.00	0.62	-0.20	-0.13	0.10	0.03	-0.04	-0.06
버스개수	0.62	1.00	-0.14	-0.01	0.30	0.23	-0.13	-0.11
*1/접근거리	-0.20	-0.14	1.00	0.13	0.10	0.08	-0.13	-0.12
*1/배차간격	-0.13	-0.01	0.13	1.00	0.34	0.62	-0.44	-0.33
운행거리	0.10	0.30	0.10	0.34	1.00	0.34	-0.39	-0.28
운행횟수	0.03	0.23	0.08	0.62	0.34	1.00	-0.39	-0.32
*1/승차인원	-0.04	-0.13	-0.13	-0.44	-0.39	-0.39	1.00	0.77
*1/하차인원	-0.06	-0.11	-0.12	-0.33	-0.28	-0.32	0.77	1.00

	지하철							
	역사개수	환승역개수	*1/접근거리	*1/배차간격	운행거리	운행횟수	*1/승차인원	*1/하차인원
역사개수	1.00	0.83	0.38	0.67	0.62	0.59	0.34	0.32
환승역개수	0.83	1.00	0.40	0.53	0.52	0.50	0.32	0.29
*1/접근거리	0.38	0.40	1.00	0.45	0.46	0.45	0.23	0.22
*1/배차간격	0.67	0.53	0.45	1.00	0.96	0.96	0.17	0.14
운행거리	0.62	0.52	0.46	0.96	1.00	0.98	0.14	0.12
운행횟수	0.59	0.50	0.45	0.96	0.98	1.00	0.08	0.05
*1/승차인원	0.34	0.32	0.23	0.17	0.14	0.08	1.00	0.99
*1/하차인원	0.32	0.29	0.22	0.14	0.12	0.05	0.99	1.0

* 변수가 커질수록 접근성 지수가 높게 산출되도록 변경

소프트파워 - 트렌드지표 3 (기업 인프라)

상관분석

상관관계가 높은 변수들끼리 결합해 최종 변수 선정

- 승차인원과 하차인원은 승하차인원으로 결합
- 승하차인원 = (승차인원 + 하차인원) / 2

단순한 정류장 개수나 역사 개수 보다는 노선의 다양성이
대중교통 접근성 판단에 더 용이하다 생각

- 정류장 개수와 버스 개수는 버스 개수를 대표 변수로 설정
- 역사 개수와 환승역 개수는 환승역 개수를 대표 변수로 설정

TCQSM에 따르면 일반적인 대중교통 서비스 평가에서
배차간격은 중요한 요소로 평가되고 있음.

- 배차간격을 제외한 운행거리, 운행횟수를 결합
 - 운행비율 = 운행횟수 / 운행거리

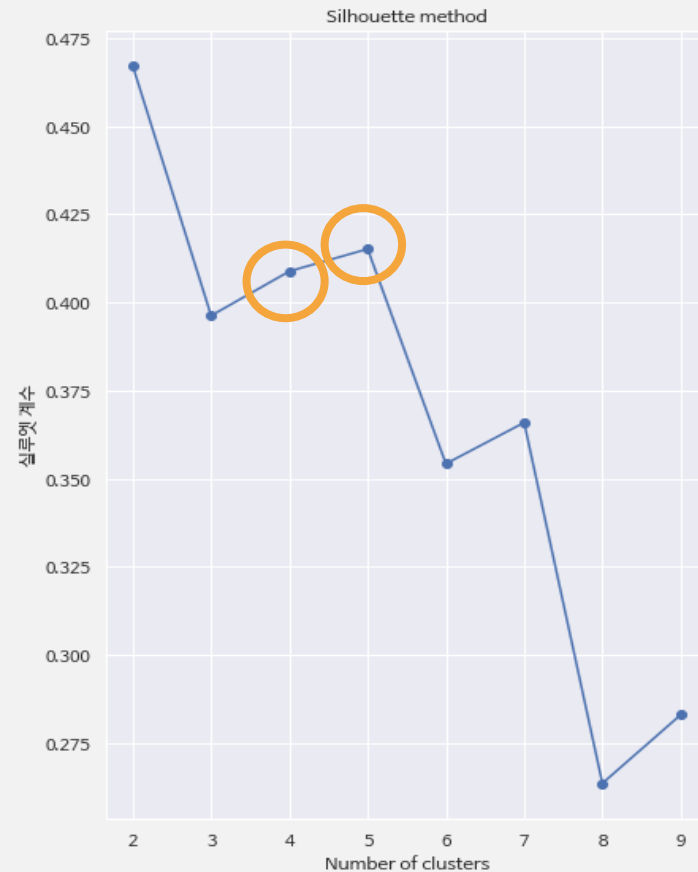
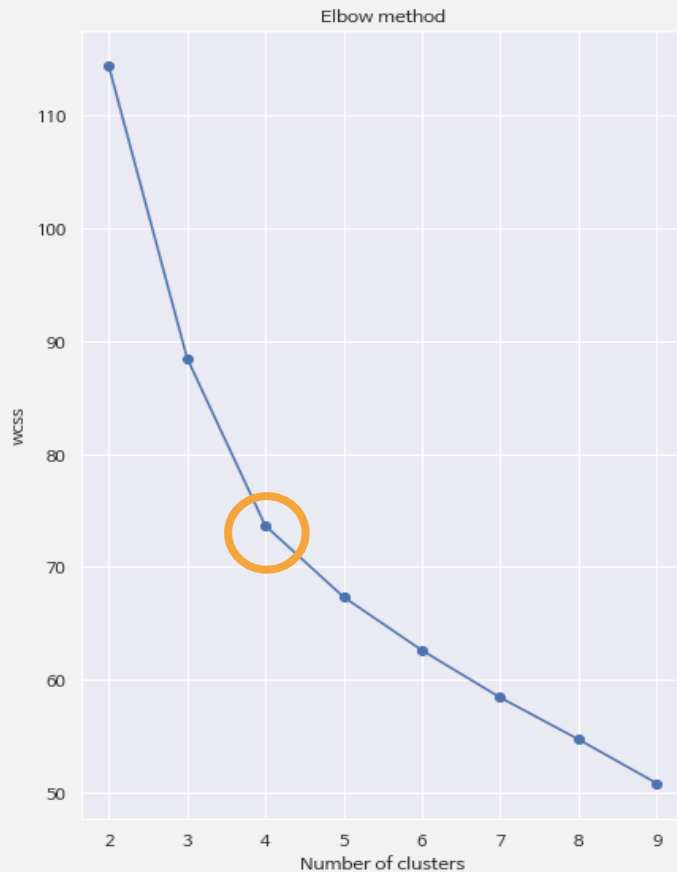
	버스									
	정류장개수	버스개수	*1/접근거리	*1/배차간격	운행거리	운행횟수	*1/승차인원	*1/하차인원		
정류장개수	1.00	0.62	-0.20	-0.13	0.10	0.03	-0.04	-0.05		
버스개수	0.62	1.00	0.13	0.19	0.03	0.23	0.12	0.06		
*1/접근거리	-0.20	0.13	1.00	0.40	0.53	-0.09	0.30	-0.11	0.17	
*1/배차간격	-0.13	0.19	0.40	1.00	0.45	-0.07	0.23	-0.06	0.06	
운행거리	0.10	0.03	0.53	0.45	1.00	-0.14	0.15	-0.13	0.50	
운행횟수	0.03	0.23	-0.09	-0.07	-0.14	1.00	0.12	0.08	-0.04	
*1/승차인원	-0.04	0.12	0.30	0.23	0.15	0.12	1.00	-0.11	-0.39	
*1/하차인원	-0.05	0.06	0.17	0.06	0.50	-0.39	0.03	1.00	0.03	
승하차인원	0.59	0.50	0.45	0.96	0.98	1.00	0.08	0.05		
*1/승하차인원	0.34	0.32	0.23	0.17	0.14	0.08	1.00	0.99		
*1/하차인원	0.32	0.29	0.22	0.14	0.12	0.05	0.99	1.0		

* 변수가 커질수록 접근성 지수가 높게 산출되도록 변경

소프트파워 - 트렌드지표 3 (기업 인프라)

상관분석

K-means 클러스터링



왼쪽 그래프 꺾이는 지점 (엘보우 포인트)의
클러스터 개수는 4개

오른쪽 그래프로 판별해본 결과
실루엣 계수가 높은 지점의 클러스터 개수는 4개, 5개

최적의 클러스터 개수를 4개로 설정

소프트파워 - 트렌드지표 3 (기업 인프라)

클러스터링

결과 해석

- 그룹 0의 경우 지하철 관련 변수는 모두 0
 - 그룹 1의 경우 평이한 값을 보임
- 그룹 2의 경우 환승역 개수, 버스 개수, 지하철 배차간격
총 3개의 변수에서 큰 값을 보임
- 그룹 3의 경우 지하철 배차간격 변수에서만 큰 값을 보임

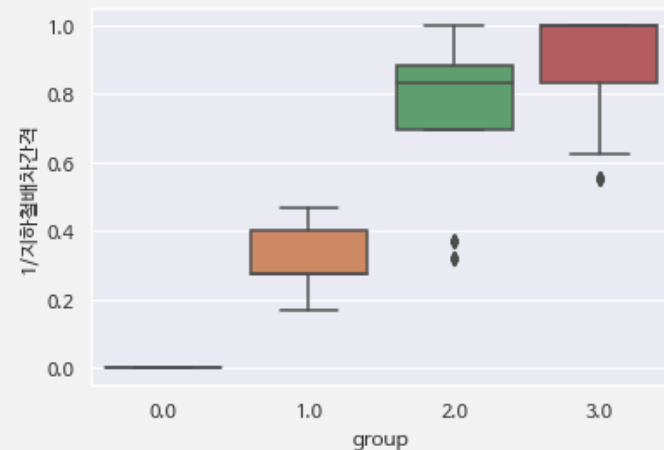
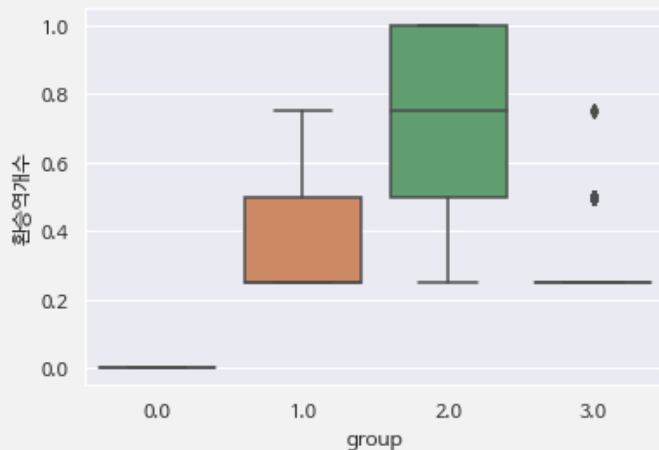
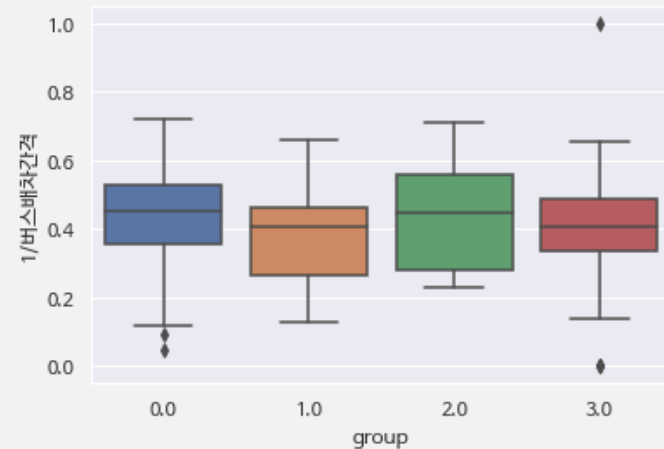
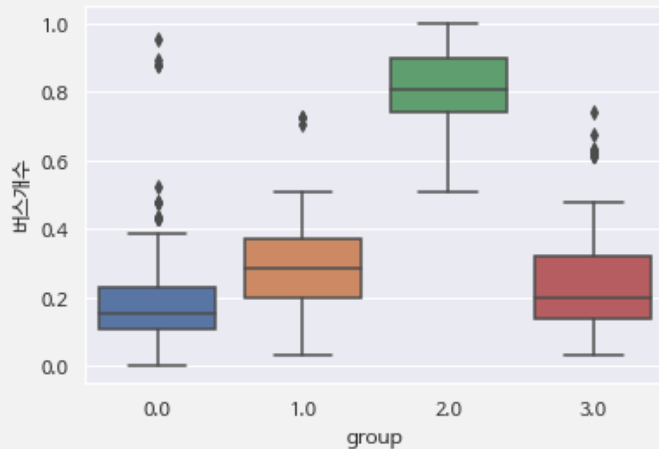
모든 변수에 동일한 가중치를 적용한다면
각 클러스터의 교통 접근성 순위는 다음과 같음

그룹 0 < 그룹 1 < 그룹 3 < 그룹 2

임의로 설정된 가중치이기에 설득력 X

>> 클러스터링을 통해 구분된 그룹 0~3을 종속변수로

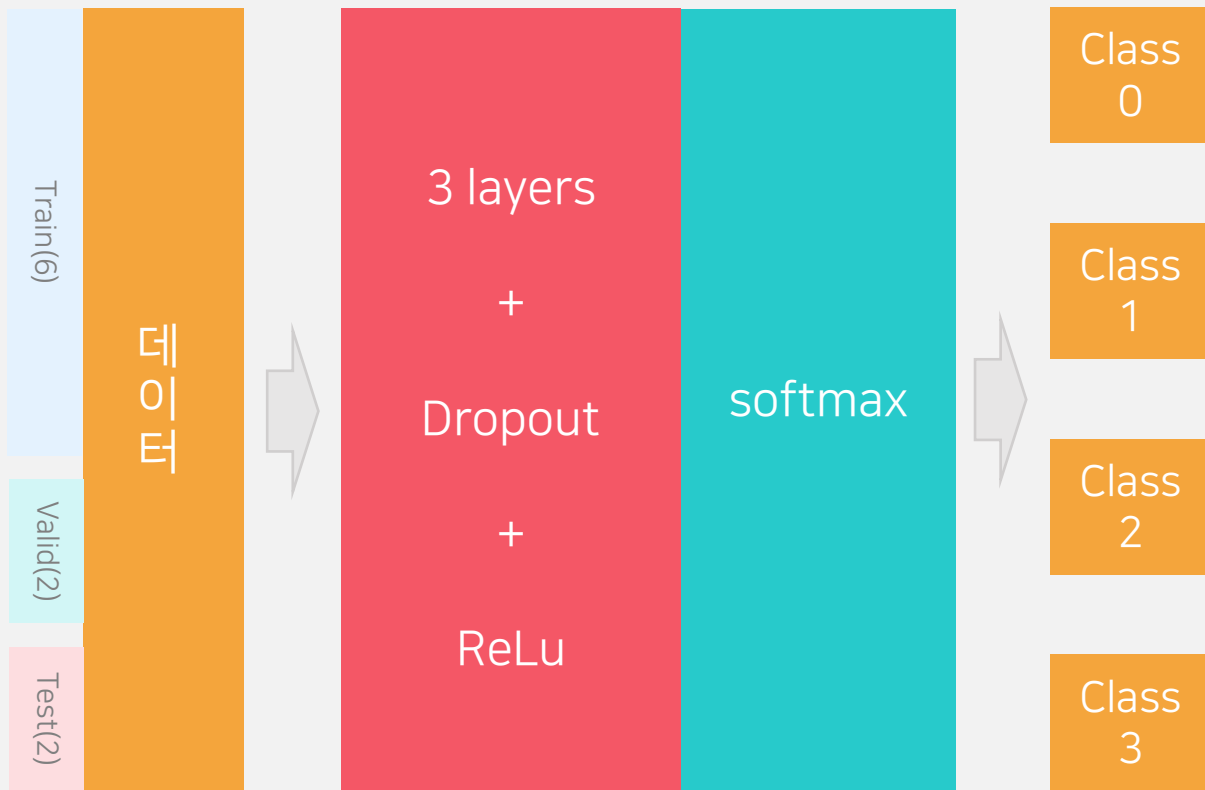
설정 후 모델링 진행



소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

딥러닝 모델(다중분류)



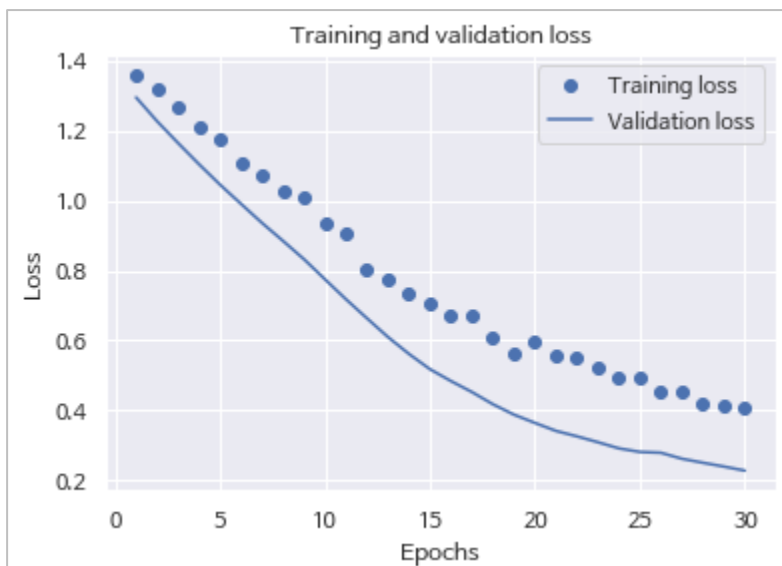
Sequential model & softmax activation function

- * 3 layers, Dropout, ReLu
- 1 Optimizer: Adam
- 2 Loss Function: categorical_crossentropy
- 3 Metric: accuracy (F1-score)
- 4 Learning Rate: 0.0005
- 5 Epochs: 30
- 6 Batch-Size: 20

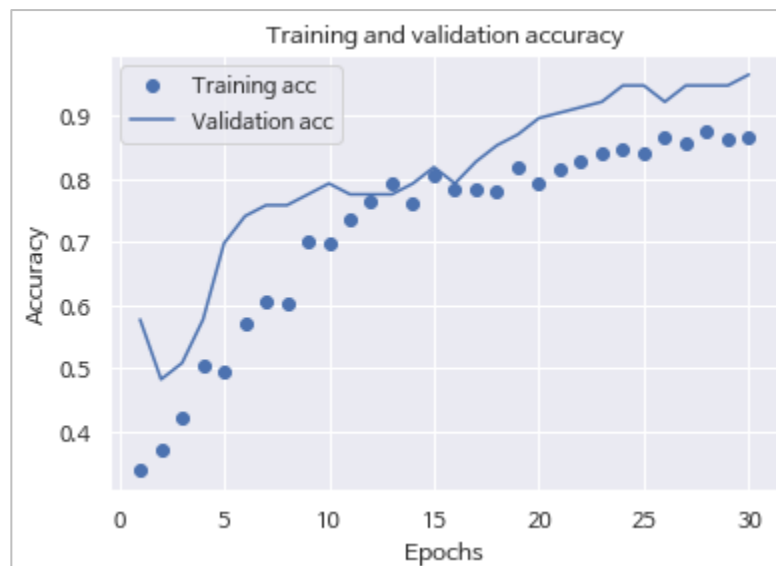
소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

딥러닝 결과 시각화



Training Loss와 Validation Loss
모두 정상적으로 줄어들과 있음



Training Accuracy가 지속적으로 높아
지고 있으므로 학습이 잘 진행되고 있음

Test Loss: 0.34
Test Accuracy: 0.89

소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

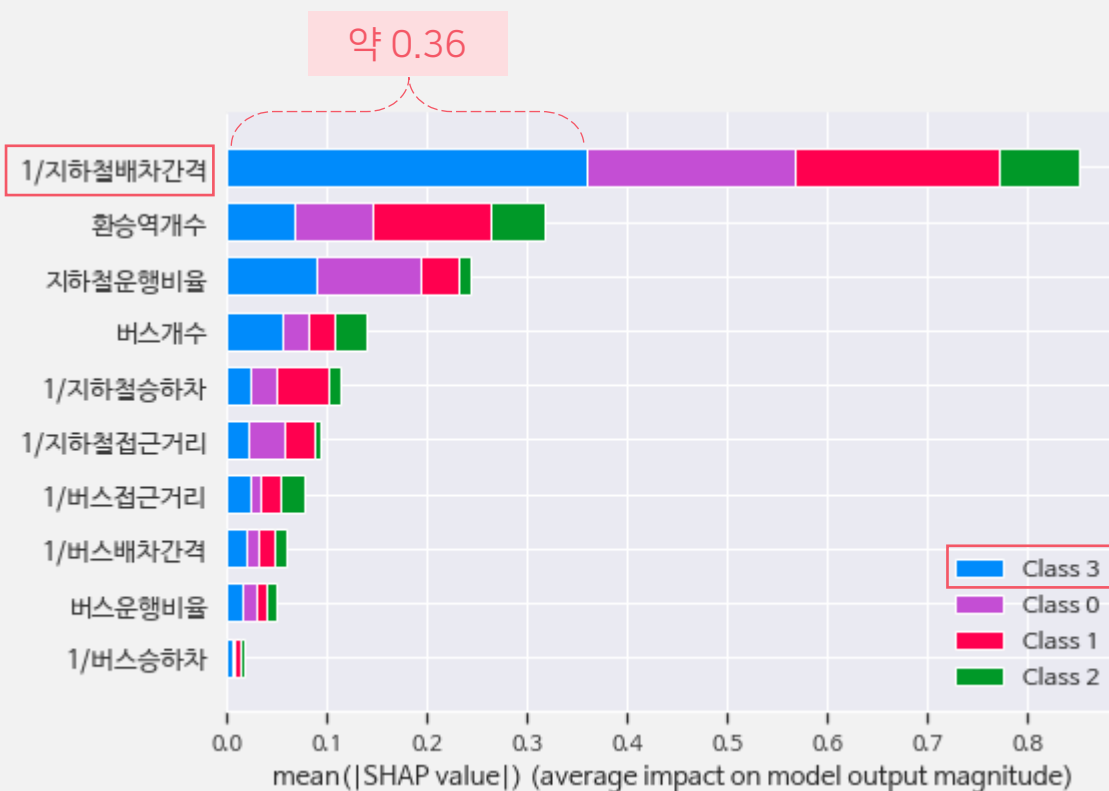
SHAP

Shapley Value의 계산 방법을 기반으로
전체적인 영역에 대한 해석 가능

목적

예측에 대한 각 특성의 기여도를 계산하여
관측치의 예측값 설명

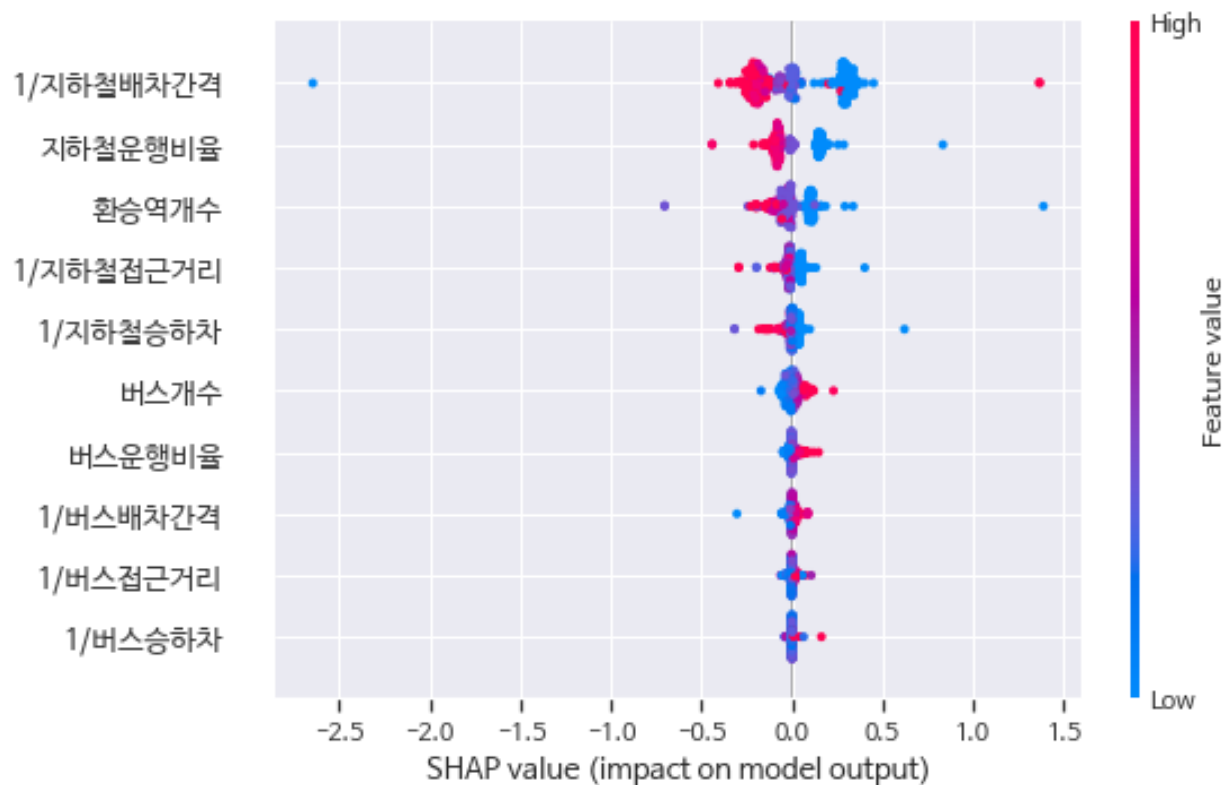
* Shapley Value: 모든 가능한 조합에 대해서 한 특성의
기여도를 종합적으로 합한 값



소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

SHAP 시각화 및 해석



Summary Plot

특성 중요도(feature importance)와
특성 효과(feature effects)를 결합

각 점은 특성에 대한 Shapley value와 관측치

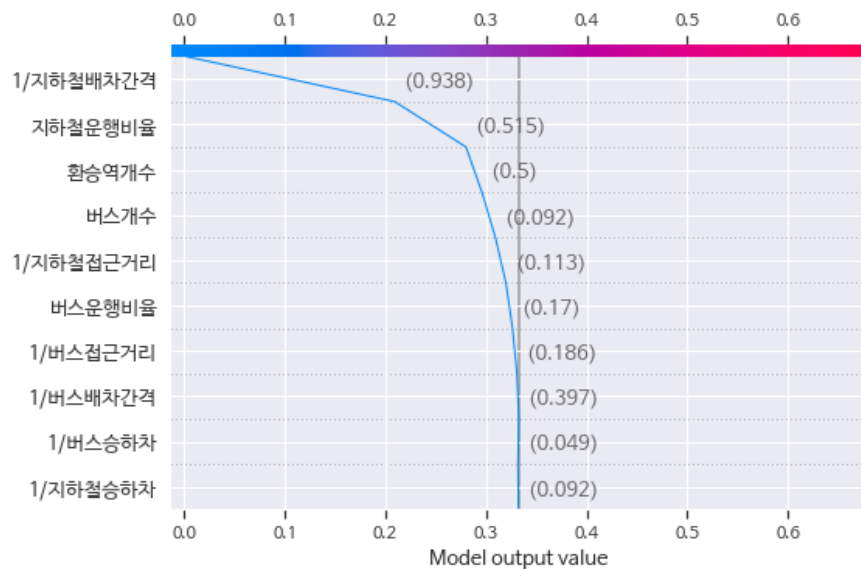
x축은 Shapley value, y축은 특성에 의해 결정

겹치는 점이 y축 방향으로 내포됨에 따라
특성 당 Shapley value의 분포 알 수 있음

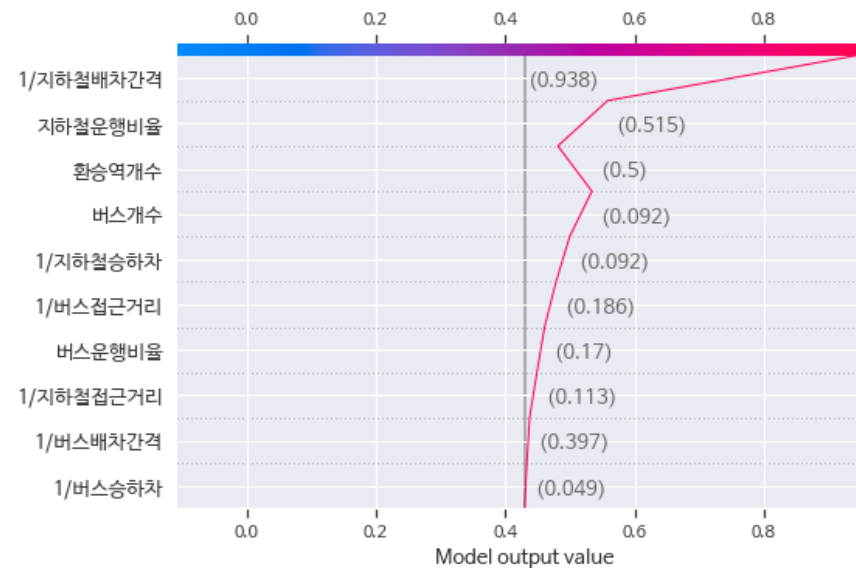
소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

SHAP 시각화 및 해석



인덱스0의 클래스0에 대한 decision plot
>> 인덱스0은 클래스0에 속하지 않음



인덱스0의 클래스3에 대한 decision plot
>> 인덱스0은 클래스3에 속함

Decision Plot

하단부터 시작한 예측선이 마지막 결과에 도착할 때까지의 누적 값을 보여줌

대략적으로 통계학 선형 모델에서 sum of effect와 같음

소프트파워 - 트렌드지표 3 (기업 인프라)

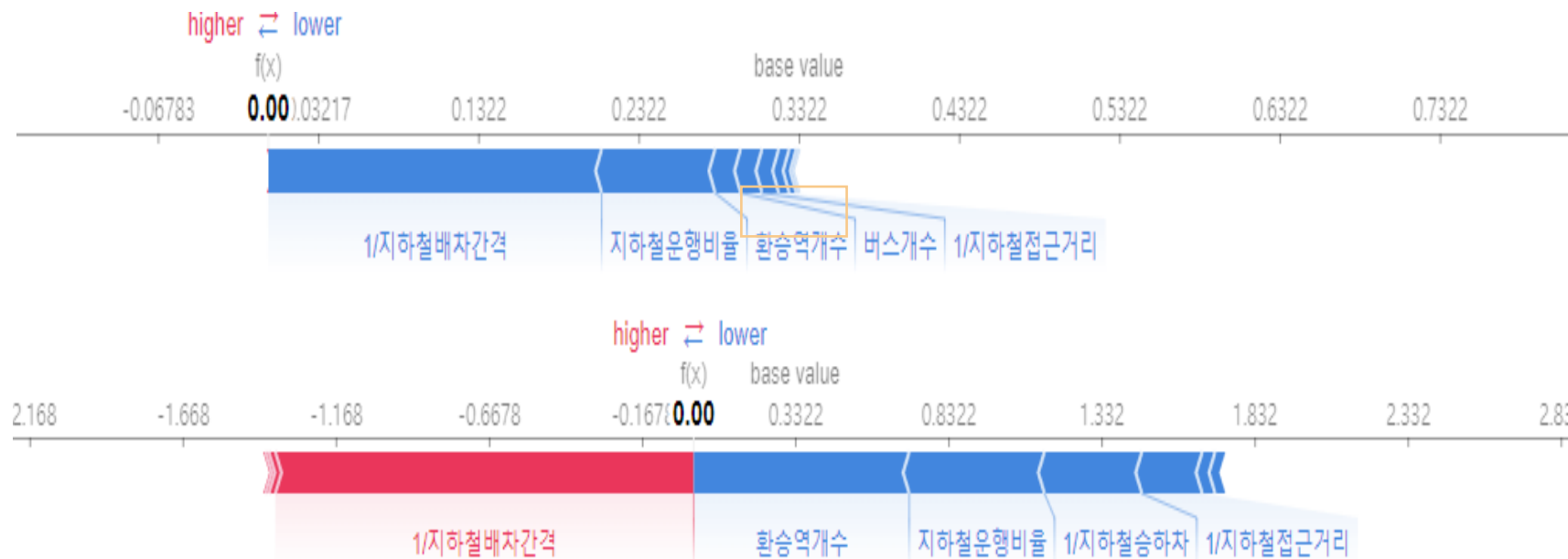
모델링

SHAP 시각화 및 해석

클래스 0일 확률을 설명하는 SHAP값,
평균 예측 확률인 기준선: 0.3322

인덱스0인 기업은 0.00의 예측확률

인덱스1인 기업도 0.00의 예측확률
But, "1/지하철배차간격"이
클래스 0으로의 예측확률에 **긍정적인** 영향



Force Plot

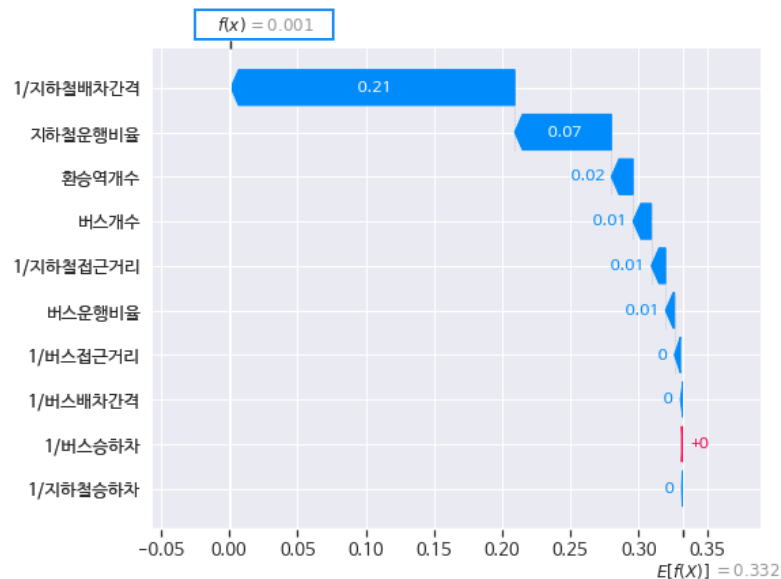
예측과 영향의 규모에 긍정적인(빨간색) 혹은 부정적인(파란색) 영향을 미치는 기능을 확인

* 클래스 0의 base value에 비해서 해당 인덱스0과 1의 output value가 어디에 위치하는지 알 수 있음

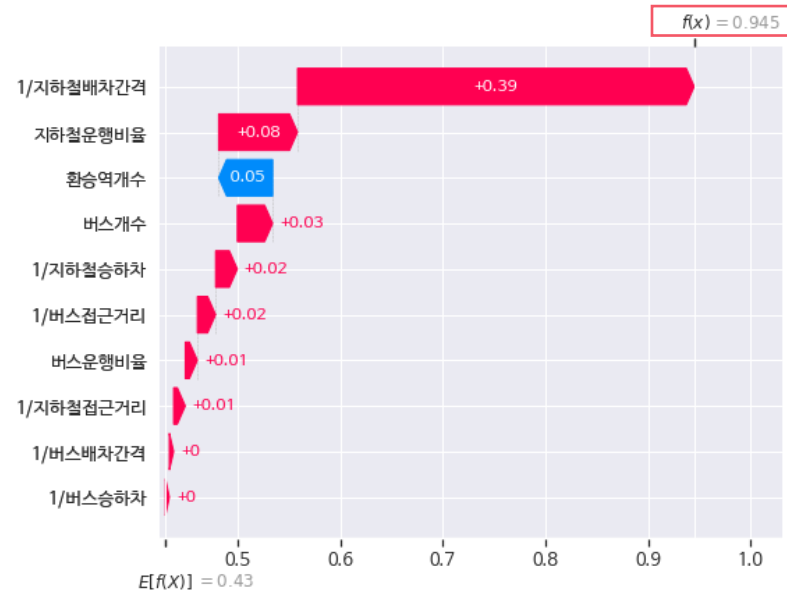
소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

SHAP 시각화 및 해석



인덱스0의 클래스0에 대한 waterfall plot
>> 인덱스0은 클래스0에 속하지 않음



인덱스0의 클래스3에 대한 waterfall plot
>> 인덱스0은 클래스3에 속함

Waterfall Plot

Force plot을 편하게 볼 수 있음

소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

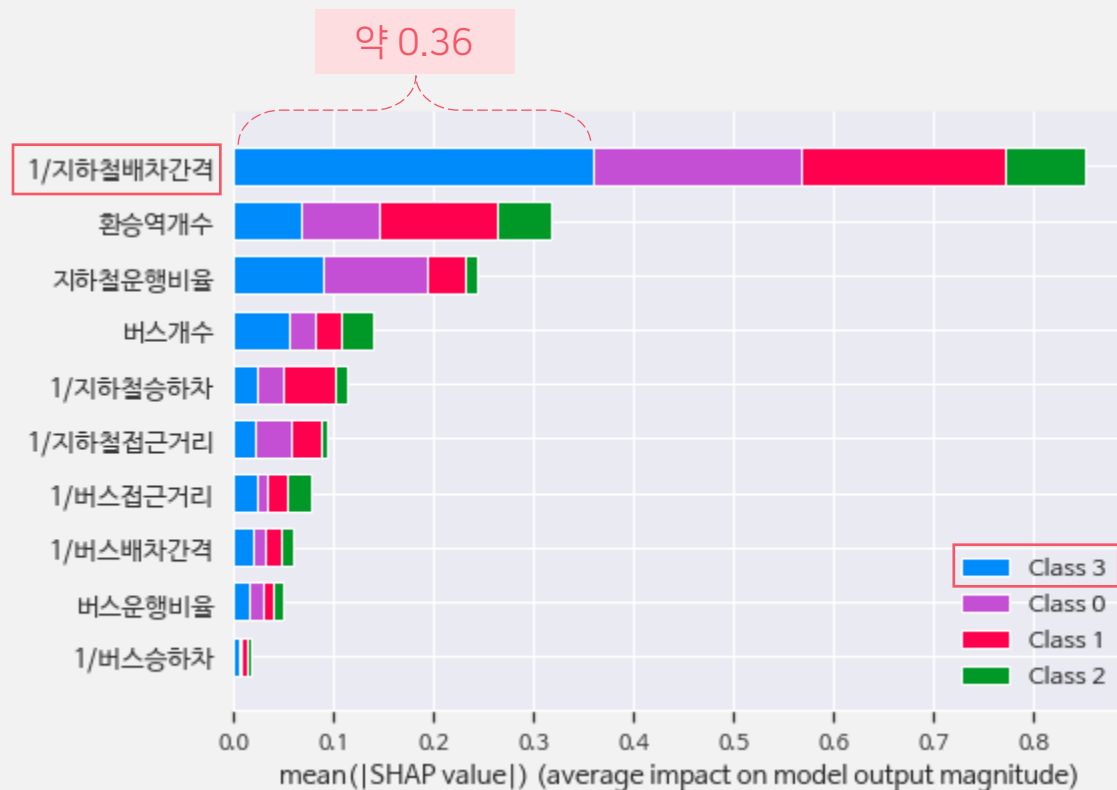
기업별 최종 접근성 지수 산출

클래스 유형	클래스별 변수 중요도 합
Class0	0.5188
Class1	0.5177
Class2	0.2470
Class3	0.6744

Mean(|SHAP Value|)는 각 클래스의 변수 중요도이므로

전체 변수들과의 비교를 통한 해당 변수의 **가중치**를

산출하기 위해서는 클래스 안에서의 변수 중요도, **비율**을 가중치로 도출



소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

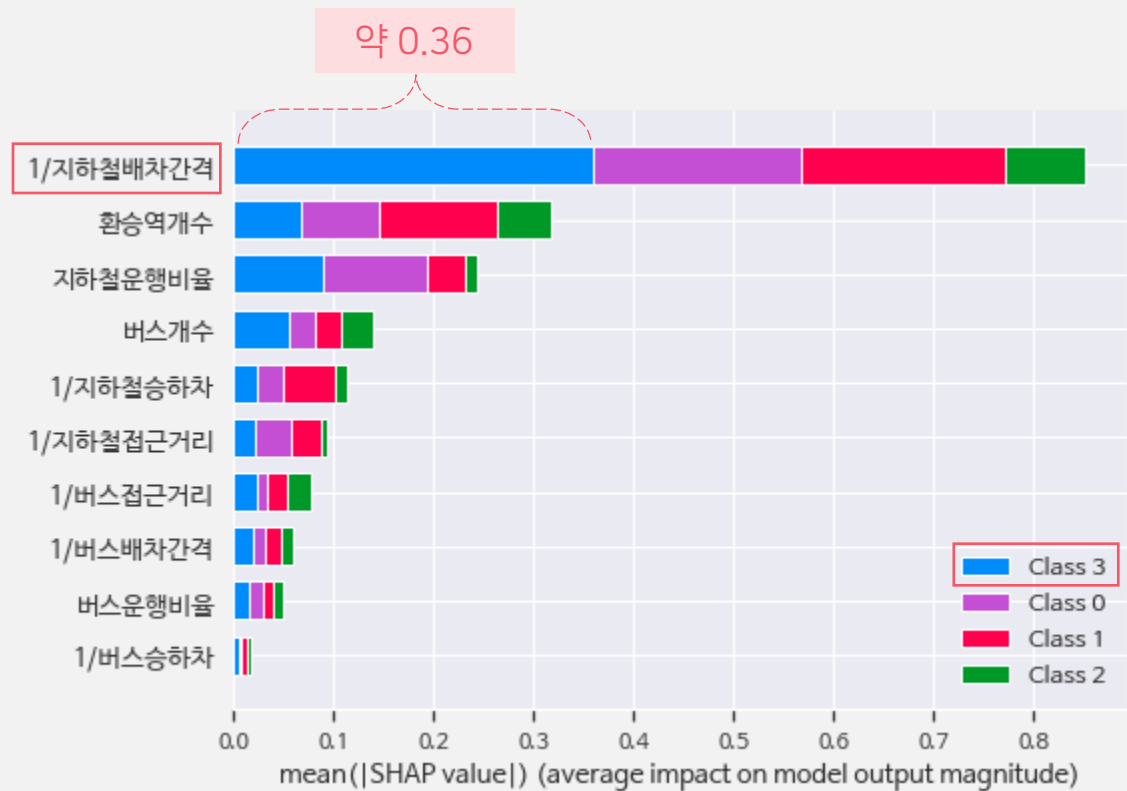
기업별 최종 접근성 지수 산출

클래스 유형	클래스별 변수 중요도 합
Class0	0.5188
Class1	0.5177
Class2	0.2470
Class3	0.6744

클래스3의 경우, Mean(|SHAP Value|)의 합은 약 0.67이므로

클래스 3의 '1/지하철 배차간격'의 가중치는

약 0.54 ($=0.36/0.67$)



소프트파워 - 트렌드지표 3 (기업 인프라)

모델링

기업별 최종 접근성 지수 산출

기업명	접근성 지수
(주)세아홀딩스	0.816767
(주)세아베스틸	0.816767
(주)세아제강지주	0.816767
애경케미칼(주)	0.532046
(주)좋은사람들	0.455627
감성코퍼레이션(주)	0.914758
SGA솔루션즈	0.009795

Min Max Scaling

접근성 지수가 가장 좋은 것은 1
가장 나쁜 것은 0으로 표기

$$\sum_{i=0}^{\text{총 변수의 개수}} \frac{\text{변수 } i \text{의 중요도}(= \text{mean}(|SHAP \text{ Value}|))}{\text{기업의 해당 클래스 변수 중요도의 합}} \times \text{해당기업 변수 } i \text{의 값}$$

Part 3

결론

MZ세대를 위한 기업 맞춤형 추천서비스

원하는 기업, 다(多) 잡(job)아봐!

[맞춤형 기업추천]

지어소프트

미투젠

아이독시

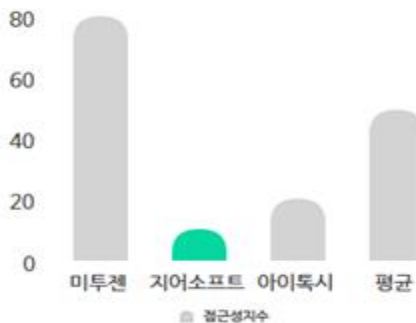
[기업 분위기 키워드]

서비스 연차 혁신
휴가 경영기술
패션 마케팅
미래 인재
교육
확보 우수 디지털

[지어소프트 레이더차트]



[접근성지수]



[지어소프트 최신이슈]

자회사인 오아시스 새벽배송 매출증가

[IT업종별 최신이슈]

-통신뉴미디어-

에그테크 투자

SK텔레콤 전자서명인증 사업자 인정

엔씨소프트의 ESG 경영

KT의 GPU기술 국산화 포부

MZ세대를 위한 기업 맞춤형 추천서비스

원하는 기업,
다(多) 잡(job)아봐!

[맞춤형 기업추천]

지어소프트

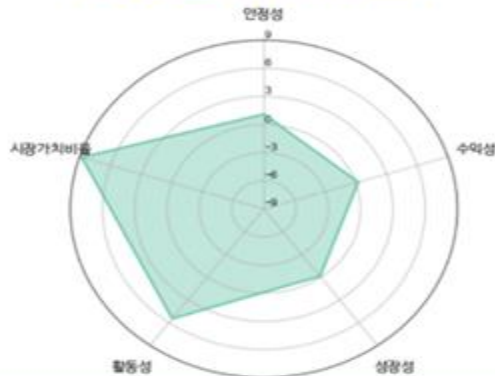
미투젠

아이독시

[기업 분위기 키워드]

출시
개발
사람
독보
소셜
시장
사랑
게임
전세계
글로벌
한국
모바일
홍콩

[미투젠 레이더차트]



[접근성지수]



[미투젠 최신이슈]

카지노 신작게임 출시와 매출 증가

자회사 미툰앤노벨의 글로벌시장 확대와
미투젠 게임 콜라보 기대

[IT업종별 최신이슈]

-인터넷-

메타버스, AI 스타트업 신규투자 확대

메타버스코인,알트코인 등 암호화폐

쏘카 차량위치안내 베타서비스 출시

MZ세대를 위한 기업 맞춤형 추천서비스

원하는 기업,
다(多) 잡(job)아봐!

[맞춤형 기업추천]

지어소프트

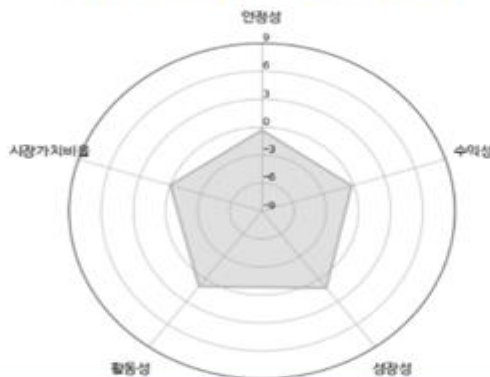
미투젠

아이독시

[기업 분위기 키워드]

국내 콘텐츠퍼블리싱
마케팅
개발
게임
전략
커머스

[아이독시 레이더차트]



[접근성지수]



[아이독시 최신이슈]

유증상자 배정으로 상장폐지 위기 극복

신작 열혈 강호 사전예약 및 이벤트 업데이트

[IT업종별 최신이슈]

-컴퓨터-

게임물관리위원회의 '무한돌파삼국지' 등급분류 취소 확정 통보

블록체인 개발사 '위메이드' 게임 사업 본격화 및 확장

각종 게임 이벤트 및 아이템,캐릭터 업데이트

MZ세대를 위한 기업 맞춤형 추천서비스

원하는 기업, 다(多) 잡(job)아봐!

【구직자 정보】

관심 직군/ 개발직군
관심 업종 / SW 개발 및 공급업

【맞춤형 기업추천】

지어소프트

미투젠

아이특시

김구직님은 수익성(26%), 안정성(13%), 활동성(3%), 성장성(50%), 시장가치비율(7%)로 기업을 평가합니다.

【미투젠 개발직군】

-합격자 평균지표-

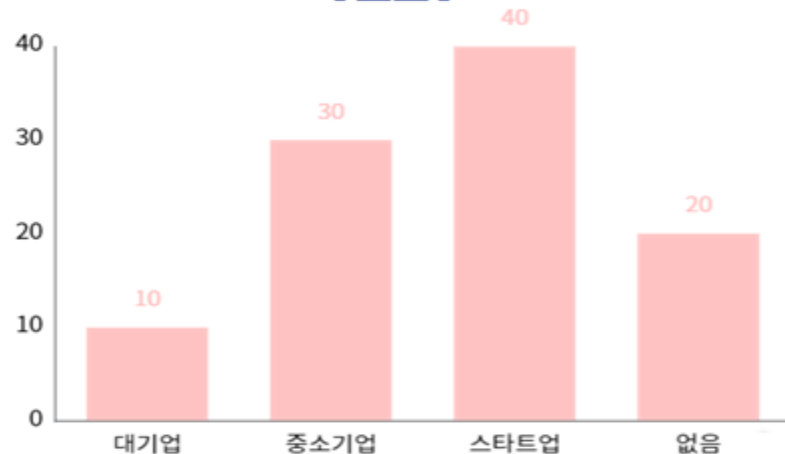
전공	컴퓨터공학과, 소프트웨어학과
학점	3.63
자격증	정보처리기사, MOS, 한국사2급
외국어	토익 830점, 토익스피킹 LV6
해외 경험	없음



【대외활동 키워드】

학회 학부연구생 특허 동아리 부트캠프
AI 인턴
기획/아이디어 논문 알고리즘 대회 해커톤

【인턴】



추가 아이디어 및 서비스

1. 마케팅 퍼포먼스

- MZ세대들에게 인기 열풍인 각종 유형의 테스트
- 이를 모티브로 재무제표 5가지 지표를 선호순으로 나타내어 해당되는 기업을 결과로 도출하는 테스트 제작
- 고용노동부 산하 워크넷 서비스에 도입하게 된다면 신규 유저 유입과 기관 홍보 효과 얻을 수 있음.



맞춤형 기업 test

재무제표로 알아보는 나에게 맞는 기업은?



선호하는 기업의 능력을 선택해주세요

- ☐ 불황에도 부도나지 않을 능력
- ☐ 매출과 이익의 성장가능성이 높은 능력
- ☒ 수익 창출로 많은 이익을 남기는 능력
- ☐ 기업의 자산을 효율적으로 이용하는 능력
- ☐ 주식 시장에서 높은 평가를 받는 능력



네이버

카카오

쿠팡

2. 기존 하드파워/소프트파워 지표 보완

- 기존 취업 정보 및 구인구직 사이트에서 제공하는 지표를 보완할 수 있는 여러 지표들 제안함.
- 고용노동부와 워크넷 측에서 다른 기관과 협업을 맺어 정보와 데이터를 확보한다면 더 큰 시너지 산출 가능함.

- 워라벨 (유연근무제, 복리후생 유무 정보)
- 임금체불 사업주 (명예훼손 문제로 링크 제공)
- 고용/산재 보험 현황 (안정성 지표)
- 청년친화강소기업 (임금, 일생활균형 부문 높은 평가)

하드파워

- ESG 기업 등급
- CEO 신년사 혹은 인터뷰 기사

소프트파워

추가 아이디어 및 서비스

3. 희망 직군/직무 맞춤서비스

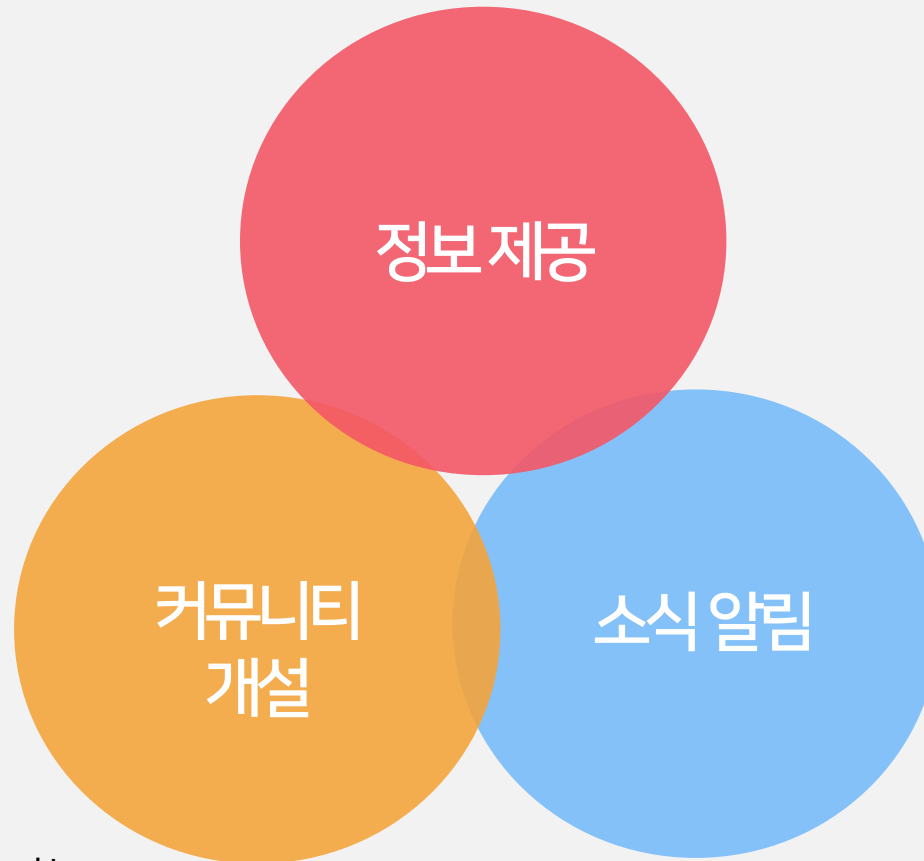
- 기업 정보만을 제공하는 것을 보완해 희망 직군/직무별로 나누어 정보 제공함.
- 기업 위주의 정보에서 직군/직무 위주의 정보 제공에 대한 필요성 체감함.

직군/직무별 커뮤니티 개설
(ex. 오픈 카카오톡, 밴드)

↓

개인별 기업 맞춤 대시보드 제공 받은 후
특정 직군/직무를 고르고 나면,
같은 분류에 묶인 타 구직자들과 소통 가능,
현직자들은 간단한 가입과 기업 리뷰를 남기면 참여 가능

사전 등록해둔 희망 직무
뉴스, 대외활동 정보 제공 서비스



키워드 등록 후 희망 직무 채용공고,
일자리 박람회 자동 알림 기능

최종 결론

의의

- 우리가 제안한 트렌드 지표는 취업률을 높일 뿐 아니라 이직률도 낮추어 기존 서비스를 보완할 수 있는 아이디어다.
- 뉴스 데이터, 해당 기업의 기업소개 페이지 등 기업분석을 할 때 구직자가 쉽게 구할 수 있는 정보 뿐 아니라, 재무제표나 해당 기업 주변 인프라 데이터 등 다양한 데이터를 사용하여 구직자가 **다각적으로 기업 분석**을 하도록 돕는다. 또한 **기업의 근본적인 변화와 더 나은 경영을 도모**할 수 있다.
- 기존 취업 정보 사이트에서 활용되지 않은 새로운 트렌드 지표를 발굴해내어 워크넷 등 정부 산하 고용 정보 기관의 새로운 **유저 유입**과 **기타 사업에 유용하게 사용할** 수 있다.
- 대기업 뿐만 아니라 구직자가 실제로 필요로 하는 (우수) 중소 기업의 정보를 사용하여 **폭 넓은 정보**를 제공한다.
- 구직자의 **개인 성향과 선호를 충분히 반영**하여 맞춤형 추천을 해준다는 점에서 뜻깊다.

한계

- 원천데이터의 양이 적은 탓에 데이터 수집하고 의미 있는 정보를 추출하는 것이 쉽지 않았다.
- 블라인드, 링커리어, 잡코리아 등 문의한 결과 사용 불가하다는 답변을 받아 활용하려 계획했던 여러 데이터를 사용할 수 없어 아이디어를 **제한적으로 구현**하게 되었다.
- 각 구직자의 거주지에서 회사까지의 경로를 고려한 접근성 지수를 도출하지 못한 것이 아쉽다.

참고문헌

- [1] '[MZ세대 열풍 리포트②] 90년대생 기업 들어왔다 떠난다...'체질개선 필요", 투데이코리아, 2021.12.22
- [2] 주지환. (2021). 재무비율의 극단치에 대한 통계적 분석. KAIST 경영대학.
- [3] 박현재. (2017). 재무비율을 이용한 기업분석 사례 연구 -아모레퍼시픽을 중심으로-
- [4] 유경상 외. (2020). 서울시 대중교통 서비스의 지역 형평성 평가. 서울연구원.
- [5] 기업경영해설. 한국은행.
- [6] 이만재 외. (2016). 고용브랜드에 대한 인식유형이 기업 구직태도에 미치는 영향에 관한 연구: 대학생 구직자의 인식을 중심으로(2016)
- [7] 노유진 외. (2007). AHP를 활용한 학술연구구성사업 발전방향에 관한 연구;이공계분야 기초연구지원을 중심으로
- [8] 김신엽 외. (2017). AHP분석을 활용한 향후 디지털 마케팅 구성요인의 중요도 연구: 부산국제광고제 애드텍 2017 사례를 중심으로

데이터 출처

- [1] 네이버 뉴스
- [2] 금융공시스템 opendart
- [3] KRX 정보데이터 시스템
- [4] 아이톡시, 지어소프트, 미투젠 홈페이지

분석도구



감사합니다