

NLP와 평점 데이터 분석 기반 기업 추천 시스템

202402489_원종찬

1. 프로젝트 개요

1.1. 무엇을 만들었나

현존하는 기업 리뷰 데이터(텍스트)와 평점(수치) 데이터를 결합하여 분석한 뒤, 사용자가 자연어 질문(예: "야근 없고 수평적인 문화인 곳")을 입력하면 그 요구사항에 가장 부합하는 기업을 찾아 추천해 주는 AI 시스템을 개발하였습니다.

1.2. 어떤 문제를 해결했나

정보의 비대칭성 및 탐색 비용 과다: 구직자가 기업의 실제 조직 문화나 분위기를 파악하기 위해 수천 개의 리뷰를 일일이 읽어야 하는 비효율적인 문제를 해결하고자 했습니다.

단순 키워드 검색의 한계: 기존 플랫폼은 '연봉', '워라밸' 등의 단순 필터링만 제공하여, "돈은 적어도 되니까 배울 게 많은 곳"과 같은 복합적이고 구체적인 니즈를 반영하기 어려웠습니다. 이를 문맥(Context) 기반 검색으로 해결했습니다.

2. 진행 과정

2.1. 주제 선정 및 문제 정의

공식적인 기업 소개 자료에서는 알기 힘든 '재직자의 진짜 목소리'를 데이터화하고자 했습니다.

단순히 긍정/부정을 분류하는 것을 넘어, "어떤 측면에서 만족하는가?"를 모델이 학습하여 사용자 질의와 매칭하는 것을 핵심 문제로 정의했습니다.

2.2. 데이터 수집 및 분석

데이터셋: 기업 리뷰 데이터(장점, 단점 텍스트)와 항목별 평점(승진 기회, 복지, 워라밸 등)을 수집하였습니다.

전처리: 특수문자 제거 및 텍스트 정제, 리뷰 텍스트와 해당 리뷰 작성자가 매긴 평점 간의 상관관계 매핑.

2.3. ML 모델 학습 및 평가

사용 모델: 한국어 처리에 특화된 Pre-trained BERT 모델 (klue/bert-base 또는 kcbert 계열 활용)

학습 전략 (Mapping Text to Score): 단순 텍스트 임베딩이 아니라, 리뷰 텍스트가 가리키는 평점(Score)을 예측하도록 모델을 학습시켰습니다. 이를 통해 모델은 "칼퇴 가능"이라는 텍스트가 "워라밸 점수 5점"과 연결된다는 의미적 관계를 학습하게 됩니다.

Loss Function: MSE(Mean Squared Error) 등을 활용하여 예측 평점과 실제 평점의 오차를 줄이는 방향으로 학습.

3. 서비스 개발

3.1. 모델을 서비스로 만든 구조

학습된 모델을 활용하여 실시간 추천이 가능한 **Inference Flow**를 구축하였습니다.

Shared Encoder: 기업 리뷰 데이터와 사용자 질문을 동일한 768차원 벡터 공간으로 투영합니다.

Vectorization: DB 내 모든 기업 리뷰를 벡터화하여 저장 (Knowledge Base) 합니다. 이 벡터에는 학습 단계에서 배운 '예상 평점 정보'가 내재되어 있습니다.

Similarity Search: 사용자 질문 벡터와 가장 코사인 유사도(Cosine Similarity)가 높은 리뷰 벡터를 탐색합니다.

4. 실제 사용 결과 (5회 이상 테스트)

개발된 시스템을 사용하여 다양한 니즈를 가진 가상의 구직자 입장에서 테스트를 진행하였습니다.

[사례 1] 워라밸 중시형

추천 결과						추천 검색
나의 선호 조건						
같이 일하는 사람들이 정상적인, 조직 문화가 좋은 곳						
추천 결과						
순위	기업명	직무	AI 적합도	실제 평점	리뷰 요약	
1	위시켓(주)	기획/경영	0.9522	5	구성원을 위한 조직문화와 복지가 좋은 회사...	
2	대그바이컴퍼니(주)	마케팅/시장조사	0.9506	4	분위기가 좋아서 회사 다니는게 자체는 힘들지않음...	
3	(주)코인원	서비스/고객지원	0.9460	4	젊은 사람들이 근무하는 수평한 느낌의 회사...	
4	콘센트릭스서비스코리아(유)	유통/무역	0.9455	3	회사가 젊으며 농지가 좋고 함께 성장하는 회사...	
5	(주)이스트소프트	개발	0.9450	4	직원들간 소통이 원활하고 대체적으로 상식이 통하는 곳...	

[사례 2] 성장 및 커리어 중심형

추천 결과						추천 검색
나의 선호 조건						
연봉이 업계 평균 이상 되고 승진 기회도 분명 확인되는 곳						
추천 결과						
순위	기업명	직무	AI 적합도	실제 평점	리뷰 요약	
1	(주)에이아이트릭스	개발	0.9362	4	인공지능 최첨단의 연구개발 능력을 접할 수 있는 회사...	
2	(주)트루엔	개발	0.9319	4	발전에 대한 가능성성이 가장 높으며 사세확장중...	
3	(주)어댑트	마케팅/시장조사	0.9232	4	더 큰 기업으로 도약을 준비 하는 곳...	
4	(주)에이프리카	전문직	0.9208	5	AI/클라우드 관련 산업이라 본인과 회사의 성장 가능성이 높은 회사...	
5	(주)팜아시아	유통/무역	0.9204	5	성장하는 회사 임직사장으로 재고시장...	

[사례 3] 급여 및 복지 실속형

나의 선호 조건

추천 검색

경영진이 젊은 비전은 가진 [개발직](#) 회사

추천 결과

순위	기업명	직무	AI 적합도	실제 평점	리뷰 요약
1	베스핀글로벌(주)	개발	0.9147	4	본인이 열심히 노력하면 분명 배울건 있음...
2	(주)반장창고	영업/제휴	0.9110	4	인테리어 분야가 어떻게 돌아가는지 경험해볼 수 있는 회사...
3	(주)奎이드	데이터	0.9051	3	스타트업에서 일하는 장점을 다 느낄수 있는곳...
4	(주)奎이드	데이터	0.9051	3	스타트업에서 일하는 장점을 다 느낄수 있는곳...
5	인테이크(주)	전문직	0.9046	4	젊은 직원 및 경영진들로 자유로운 분위기...

[사례 4] 수평적 조직문화 선호형

나의 선호 조건

추천 검색

출퇴근이 자유롭고 연봉도 높은 곳

추천 결과

순위	기업명	직무	AI 적합도	실제 평점	리뷰 요약
1	(주)와이즈넷	데이터	0.9311	5	자유롭고 수평적임 첫회사로 다니기 좋음...
2	(주)리브스메드	전문직	0.9246	4	디플 열심히 으쌰으쌰하는 분위기에 긍정적인 분위기인 것 같습니다...
3	(주)에프엔에프	인사/총무	0.9241	5	기업 규모에 비해 매우 자유로운 분위기의 회사...
4	(주)데일리펀딩	마케팅/시장조사	0.9239	5	개인의 성장 가능성은 무궁무진. 단 업권 규제는 넘기 힘든 산...
5	콘센트릭스서비스코리아(유)	서비스/고객지원	0.9230	5	외국계 기업이고 복지 및 균무시설좋음...

[사례 5] 구체적 복지 선호형

나의 선호 조건

추천 검색

개발직 직무를 배우고 성장할 수 있는 곳. 연봉은 큰 상관 없음.

추천 결과

순위	기업명	직무	AI 적합도	실제 평점	리뷰 요약
1	(주)미리디	기획/경영	0.9171	4	신입으로 입사하여 성장할수 있었습니...
2	(주)팜아시아	영업/제휴	0.9102	4	성장의 발판이 되어도 좋은 곳.....
3	(주)드림어스컴퍼니	디자인	0.9087	5	디자인 전체 프로세스를 경험할 수 있는 곳...
4	(주)플래티어	개발	0.9079	4	신입으로 들어와서 경력쌓고 발판으로 삼아서 이직하기에 괜찮은회사...
5	(주)쥬먼컨설팅그룹	개발	0.9073	4	기업문화가 나쁘지 않은곳으로 배우고 성장할 수 있는 곳이다...

5. 배운 점 및 개선 방향

5.1. 배운 점

텍스트와 수치 데이터의 결합: 단순히 텍스트만 분석하는 것보다, 해당 텍스트가 실제로 어떤 평점(만족도)과 연결되는지 함께 학습시켰을 때 추천의 정확도가 비약적으로 상승함을 확인했습니다.

벡터 검색의 유용성: 벡터 공간 내 유사성을 통해 의미론적 검색이 가능함을 경험했습니다.

5.2. 한계점 및 개선 방향

데이터의 최신성: 크롤링 시점의 데이터에 의존하므로, 최근 기업 이슈가 즉각 반영되지 않을 수 있습니다. 주기적인 데이터 파이프라인 구축이 필요합니다.

필터링 기능 추가: 현재는 의미 기반 추천만 가능하지만, 실제 서비스화를 위해서는 '지역(서울/판교 등)', '산업군', '기업 규모' 등의 메타데이터 필터를 결합한 하이브리드 추천 시스템으로 발전시킬 필요가 있습니다.

속도 최적화: 전체 리뷰 데이터를 매번 검색하는 방식은 데이터가 커질수록 느려질 수 있으므로, Faiss와 같은 고속 벡터 검색 라이브러리 도입을 고려해야 합니다.