

인공지능 기반 커리어 매칭 시스템

캡스톤 디자인 25팀

목차

- 팀원 소개
- 프로젝트 소개
- 진행내용
- 예정사항

팀 소개

- 이호승

- ID: 20163149
- E: lhs971106@kookmin.ac.kr
- Folder: <https://github.com/stigr>

- 박정재

- ID: 20150359
- E: vcz1017@kookmin.ac.kr
- Folder: <https://github.com/3ltigr0>

회사 소개

- 그렙

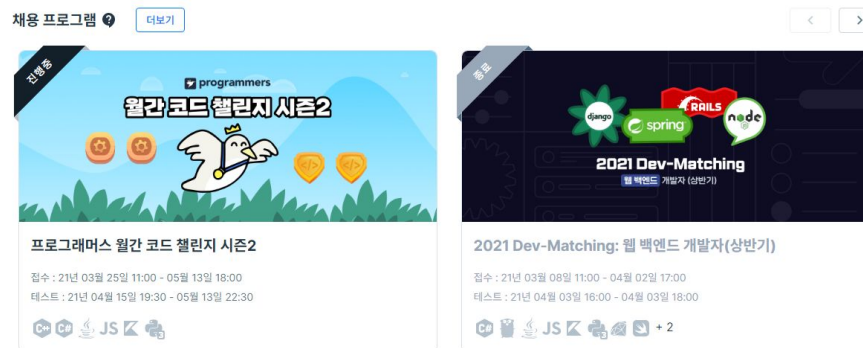
- Grepp, Inc.는 개발자들의 경쟁력 강화를 위한 교육, 교류, 문제 해결을 위한 모든 활동을 지원하며, 누구나 소프트웨어 개발에 참여할 수 있도록 도움을 주고 있습니다.



프로젝트 소개

프로그래머스

- 개발자를 위한 프로그래밍 학습, 알고리즘 문제 풀이, 코딩 테스트, 개발자 채용 등을 제공하는 플랫폼



개발자 채용

943개의 포지션

최신순 ▼



플랫폼 개발

게임원 평균 6일 이내 응답

🏠 경력 무관 📍 서울시 송파구 💰 사이닝 보너스 500,000

서버/백엔드 프론트엔드 웹 풀스택 외 3개



Back-end Developer

미소 평균 2주 이내 응답

🏠 3 ~ 15년 경력 📍 서울 서초구 💰 사이닝 보너스 500,000

서버/백엔드 웹 풀스택 JavaScript 외 5개



웹 백엔드 개발자

더브이씨 평균 2일 이내 응답

🏠 3 ~ 10년 경력 📍 서울 마포구 💰 사이닝 보너스 500,000

서버/백엔드 웹 풀스택 ExpressJS 외 9개




웹 프론트엔드 개발자

비즈니스캔버스 평균 3일 이내 응답

🏠 1 ~ 15년 경력 📍 서울 강남구 💰 사이닝 보너스 500,000

프론트엔드 웹 풀스택 CSS 외 5개



백엔드 개발자

비즈니스캔버스 평균 3일 이내 응답

🏠 1 ~ 15년 경력 📍 서울 강남구 💰 사이닝 보너스 500,000

서버/백엔드 웹 풀스택 Node.js 외 5개



주니어 백엔드 개발자

솔라커넥트 평균 1주 이내 응답

🏠 경력 무관 📍 서울 강남구 💰 사이닝 보너스 500,000

서버/백엔드 웹 풀스택 데이터 엔지니어 외 5개

<p>플랫폼 개발</p> <p>GAMEFJNE 게임원 평균 4일 이내 응답</p> <p>경력 무관 서울시 송파구 사이닝 보너스 500,000</p> <p>서버/백엔드 프론트엔드 웹 플스택 외 3개</p>	<p>Back-end Developer</p> <p>miso 미소 평균 2주 이내 응답</p> <p>3 ~ 15년 경력 서울 서초구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 JavaScript 외 5개</p>
<p>웹 백엔드 개발자</p> <p>다브이씨 THEVC 평균 2일 이내 응답</p> <p>3 ~ 10년 경력 서울 마포구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 ExpressJS 외 9개</p>	<p>웹 프론트엔드 개발자</p> <p>BUSINESS CANVAS 비즈니스캔버스 평균 3일 이내 응답</p> <p>1 ~ 15년 경력 서울 강남구 사이닝 보너스 500,000</p> <p>프론트엔드 웹 플스택 CSS 외 5개</p>
<p>백엔드 개발자</p> <p>BUSINESS CANVAS 비즈니스캔버스 평균 3일 이내 응답</p> <p>1 ~ 15년 경력 서울 강남구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 Node.js 외 5개</p>	<p>주니어 백엔드 개발자</p> <p>솔라커넥트 평균 1주 이내 응답</p> <p>경력 무관 서울 강남구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 데이터 엔지니어 외 5개</p>
<p>Commerce 솔루션 백엔드 엔지니어(Back-end...)</p> <p>M X N COMMERCE 영역스오텍마스코리아 평균 4일 이내 응답</p> <p>5 ~ 15년 경력 서울 성동구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 시스템 소프트웨어 외 5개</p>	<p>Machine Learning Engineer</p> <p>TotemManager 테이플매니저</p> <p>경력 무관 서울 성동구 사이닝 보너스 500,000</p> <p>머신러닝 인공지능(AI) 데이터 엔지니어 외 5개</p>
<p>Back-end Developer</p> <p>TotemManager 테이플매니저</p> <p>3 ~ 7년 경력 서울 성동구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 Node.js 외 1개</p>	<p>챗봇 솔루션 개발 연구원</p> <p>미니소프트 평균 2주 이내 응답</p> <p>3 ~ 10년 경력 서울 강남구 사이닝 보너스 500,000</p> <p>서버/백엔드 프론트엔드 응용 프로그래밍 외 5개</p>
<p>웹개발자</p> <p>미니소프트 평균 2주 이내 응답</p> <p>3 ~ 10년 경력 서울 강남구 사이닝 보너스 500,000</p> <p>서버/백엔드 프론트엔드 응용 프로그래밍 외 5개</p>	<p>백엔드 개발자</p> <p>DEARGEN 디어젠</p> <p>3 ~ 5년 경력 서울 서초구 사이닝 보너스 500,000</p> <p>서버/백엔드 인공지능(AI) Redis 외 4개</p>
<p>프론트엔드 개발</p> <p>DEARGEN 디어젠</p> <p>0 ~ 3년 경력 서울 서초구 사이닝 보너스 500,000</p> <p>프론트엔드 인공지능(AI) JavaScript 외 3개</p>	<p>Back-end 개발자</p> <p>iPortfolio iPortfolio 평균 5일 이내 응답</p> <p>경력 무관 서울특별시 중구 사이닝 보너스 500,000</p> <p>서버/백엔드 웹 플스택 Spring 외 4개</p>
<p>Front-end 개발자</p> <p>iPortfolio iPortfolio 평균 5일 이내 응답</p> <p>경력 무관 서울특별시 중구 사이닝 보너스 500,000</p> <p>서버/백엔드 프론트엔드 웹 플스택 외 7개</p>	<p>Application Engineer - Tech Lead</p> <p>SOICE- 세이지리서치주식회사</p> <p>3 ~ 8년 경력 서울 서초구 사이닝 보너스 500,000</p> <p>서버/백엔드 시스템/네트워킹 응용 프로그래밍 외 2개</p>
<p>모바일 앱 엔지니어</p> <p>RIDI 리디 평균 1주 이내 응답</p> <p>4 ~ 10년 경력 서울특별시 강남구 사이닝 보너스 500,000</p> <p>안드로이드 앱 아이폰 앱 React Native 외 2개</p>	<p>[나우웨이팅] 안드로이드 엔지니어</p> <p>나우비스킹 평균 1주 이내 응답</p> <p>3 ~ 12년 경력 서울 강남구 사이닝 보너스 500,000</p> <p>안드로이드 앱 Kotlin Android 외 1개</p>
<p>QA</p> <p>TAGby TAGby 평균 4일 이내 응답</p>	<p>Web Developer(Python)</p> <p>TAGby TAGby 평균 4일 이내 응답</p>

943개의 포지션

너무 많다

나와 연관순으로 보기

- 사용자의 프로필에 기반해 적합한 순서로 목록 출력

직무 ▼

경력 ▼

연봉 ▼

직원 수 ▼

지역 ▼

942개의 포지션

현재 조건의 새로운 포지션 알림 ☐ [포지션 알림 목록 →](#)

나와 연관순 ▼

프로필을 등록하고 추천 포지션을 확인하세요!
개발자 맞춤형 프로필을 등록하면,
홍길동님과 기술 궁합이 잘 맞는 포지션을 추천해드립니다.
[프로필 등록](#) [내 이력서로 간편 등록](#)

Back-end Developer
미소 평균 2주 이내 응답
3 ~ 15년 경력 서울 서초구 사이닝 보너스 500,000
서버/백엔드 웹 풀스택 Ruby on Rails 외 5개

웹 백엔드 개발자
더브이씨 평균 2일 이내 응답
3 ~ 10년 경력 서울 마포구 사이닝 보너스 500,000
서버/백엔드 웹 풀스택 ES6 외 9개

웹 프론트엔드 개발자
비즈니스캔버스 평균 3일 이내 응답
1 ~ 15년 경력 서울 강남구 사이닝 보너스 500,000
프론트엔드 웹 풀스택 Webpack 외 5개

프로필 등록

- ☐ 경력
- ☒ 개인 프로젝트
- ☒ 주요 기술
- ☐ 외국어
- ☐ 학력
- ☐ 활동
- ☐ 수상 경력/자격증
- ☐ 출판/논문/특허
- ☐ 챌린지/과제관 결과
- ☐ 첨부파일
- ☒ 저장소
- ☐ 제외 기업

- ☒ 직군 *
- ☒ 연봉 *
- ☒ 연락처 *

다음 중 자신 있는 기술은 무엇인가요? (최대 3개)

☒ Pandas

☒ C++

☒ PyTorch

더 많은 태그 추가

희망하는 직군을 선택하세요. (최대 3개)

☐ 서버/백엔드

☐ 프론트엔드

☐ 웹 풀스택

☐ 안드로이드 앱

☐ 아이폰 앱

☒ 머신러닝

☒ 인공지능(AI)

☒ 데이터 엔지니어

☐ 모바일 게임

☐ 게임 클라이언트

☐ 게임 서버

☐ 시스템/네트워크

☐ 시스템 소프트웨어

☐ 인터넷 보안

☐ QA

☐ 사물인터넷(IoT)

☐ 응용 프로그램

☐ 블록체인

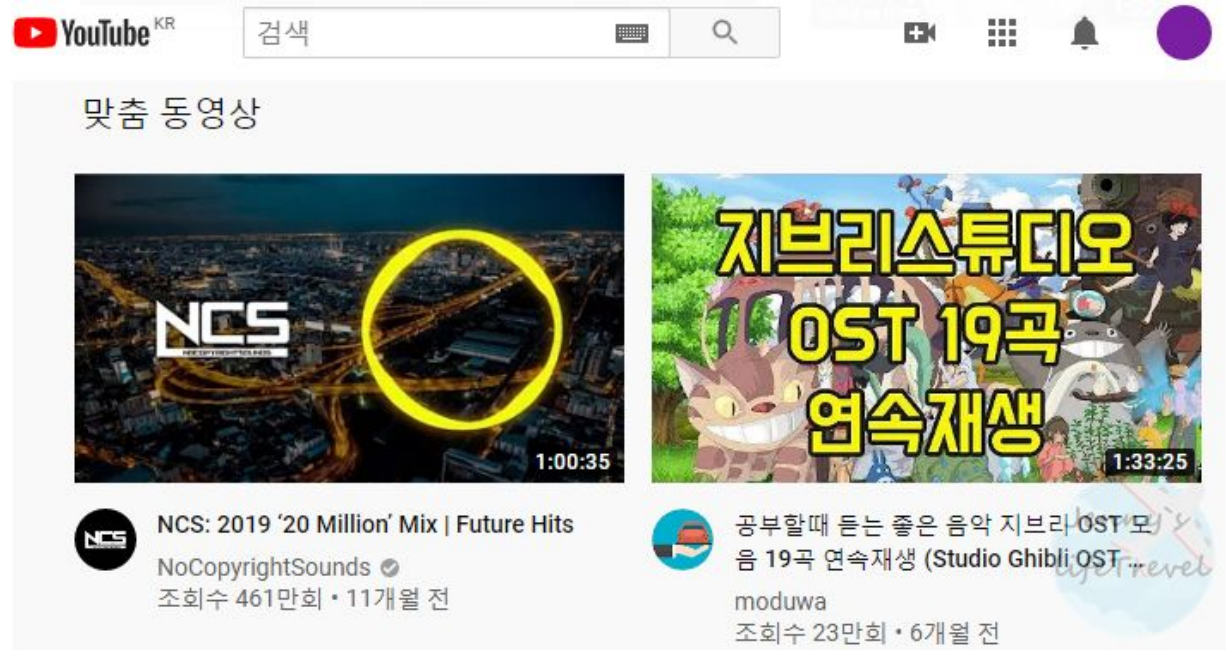
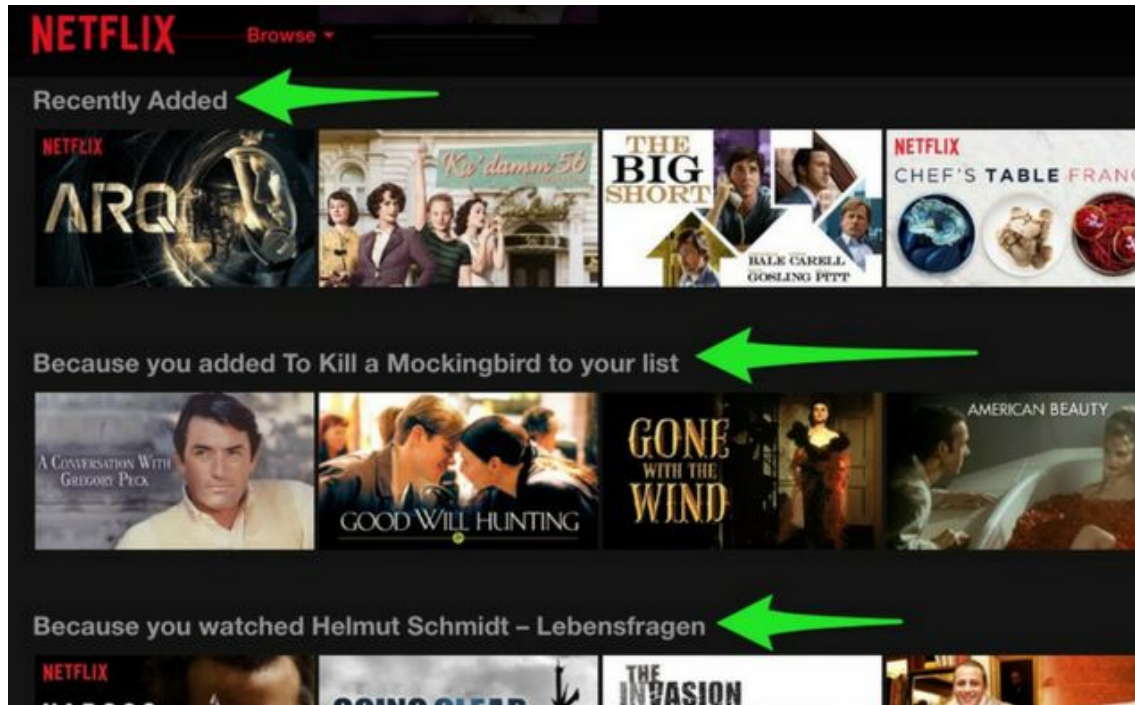
나와 연관순으로 보기

- 프로필 수정, 공고 조회 등 여러 활동을 해보았는데 큰 변화가 없음.



개인화 추천 시스템

- 사용자의 다양한 명시, 암시적 데이터 반영
 - 클릭 기록, 관심사 등



프로젝트 소개

● 인공지능 기반 커리어 매칭 시스템

- 그랩이 보유하고 있는 데이터를 기반으로 개인별 스킬셋과 경험에 맞는 커리어를 매칭해 주는 인공지능 시스템을 고안하고 프로토타입을 개발한다.

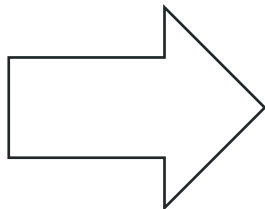



Personal
data
지원 기록
클릭 기록

프로필 등록

- ✓ 경력
- ✓ 개인 프로젝트
- ✓ 주요 기술
- ✓ 외국어
- ✓ 학력
- ✓ 활동
- ✓ 수상 경력/자격증
- ✓ 출판/논문/특허
- ✓ 챌린지/과제관 결과
- ✓ 첨부파일
- ✓ 저장소
- ✓ 제외 기업

- ✓ 직군 *
- ✓ 연봉 *
- ✓ 연락처 *





시스템 프로그램 C, Shell, IPC 등 / 대외접속...

한국자금중개 **평균 3주 이내 응답**

3 ~ 7년 경력 | 서울특별시 종로구 | 사이닝 보너스 500,000

서버/백엔드 | 시스템/네트워크 | 시스템 소프트웨어 | 외 4개



백엔드 개발자(B2C 풀리).

프리윌린

2 ~ 5년 경력 | 서울시 관악구 | 사이닝 보너스 500,000

서버/백엔드 | Spring | Kotlin | 외 1개



iOS 개발자(B2C 풀리).

프리윌린

3 ~ 8년 경력 | 서울시 관악구 | 사이닝 보너스 500,000

아이폰 앱 | Swift | iOS | 외 1개

진행 내용

진행 내용

- 그립 데이터를 받기 전 관련 기술 이해
- 데이터에 적합한 모델 탐색
- 모델에 맞게 데이터 / 평가방법 수정

진행 내용

일정	1/22	1/29	2/5	2/12	2/19	2/26	3/5	3/12	3/19	3/26	4/2
인공지능 스터디											
샘플 데이터 확보											
샘플 데이터 분석											
관련기술 분석											
기술 적용											

Toy Project 진행

- 프로그래머스 과제관의 [머신러닝] 채용 공고 추천
 - 프로그래머스 채용 공고 페이지를 방문한 개발자들의 방문/지원 기록을 바탕으로 추천 모델을 만드는 것
 - 개발자와 채용 공고를 보고, 개발자가 해당 채용 공고에 지원할지 안 할지를 예측하는 이진 분류기를 제작하는 것

데이터

- 프로그래머스 데이터 사용
 - User가 job에 지원한 정보 데이터 활용



user1

지원한 구인공고

시스템 프로그램 C, Shell, IPC 등 / 대외접속...

한국자금융중개 Korea Money Brokerage Corp.

한국자금융중개 **평균 3주 이내 응답**

🏢 3 ~ 7년 경력 📍 서울특별시 종로구 🎁 사이닝 보너스 500,000

서버/백엔드 시스템/네트워크 시스템 소프트웨어 외 4개

확인했지만 지원하지 않은 구인공고

백엔드 개발자(B2C 풀리).

freewheelin 프리윌린

🏢 2 ~ 5년 경력 📍 서울시 관악구 🎁 사이닝 보너스 500,000

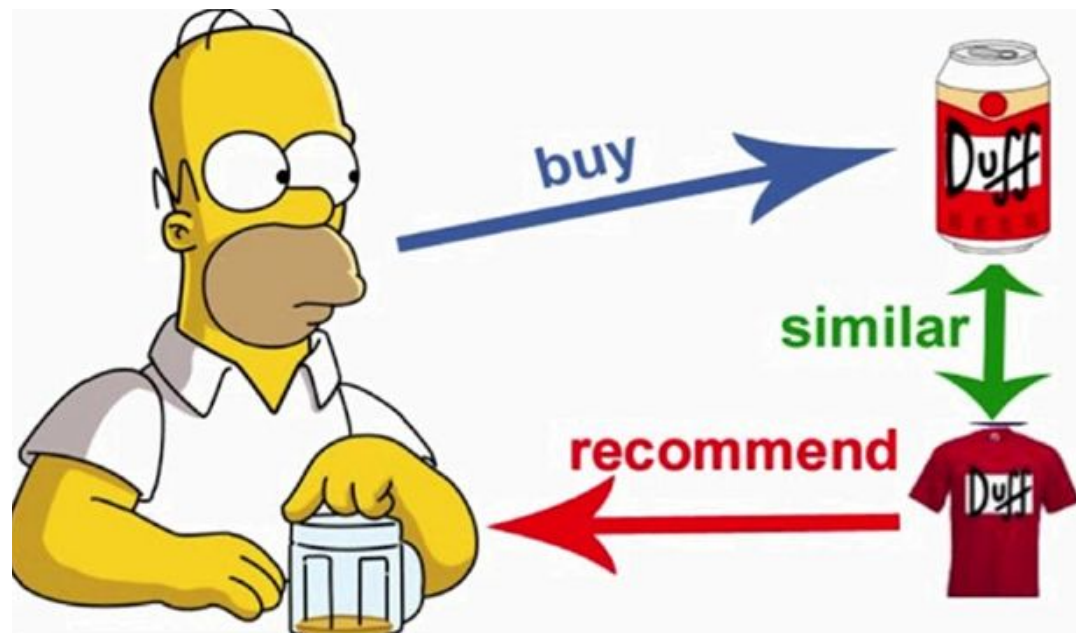
서버/백엔드 Spring Kotlin 외 1개

추천 시스템

Collaborative filtering



Content based filtering



NCF 모델 적용

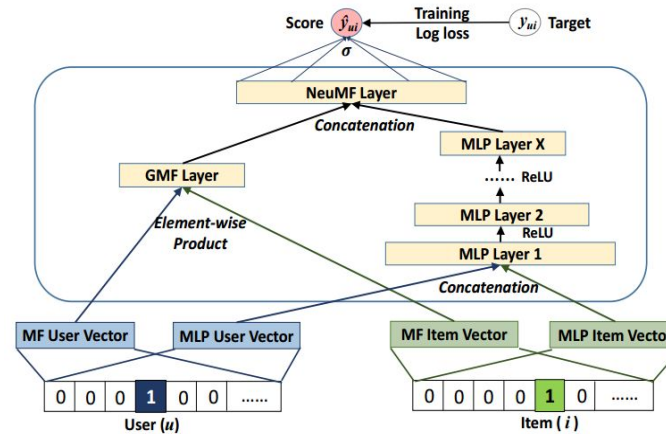
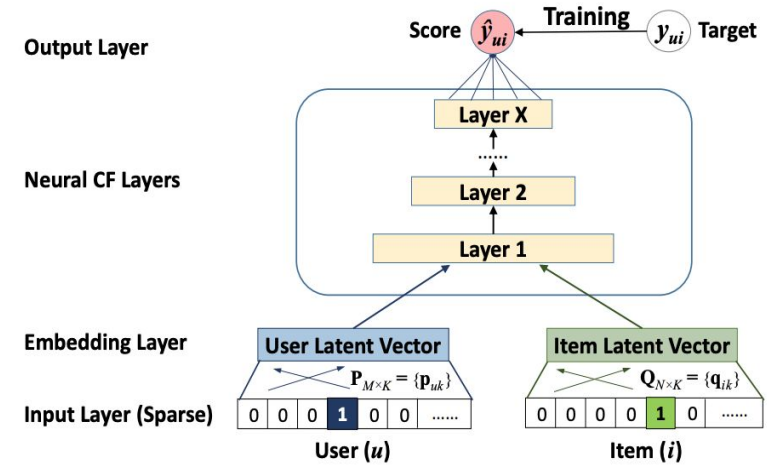
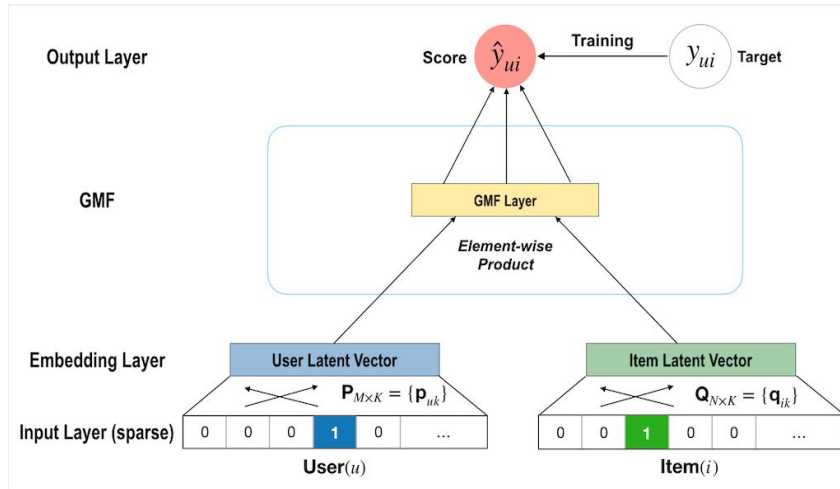


Figure 3: Neural matrix factorization model

Implicit Data

				
John 	5	1	3	5
Tom 	?	?	?	2
Alice 	4	?	3	?

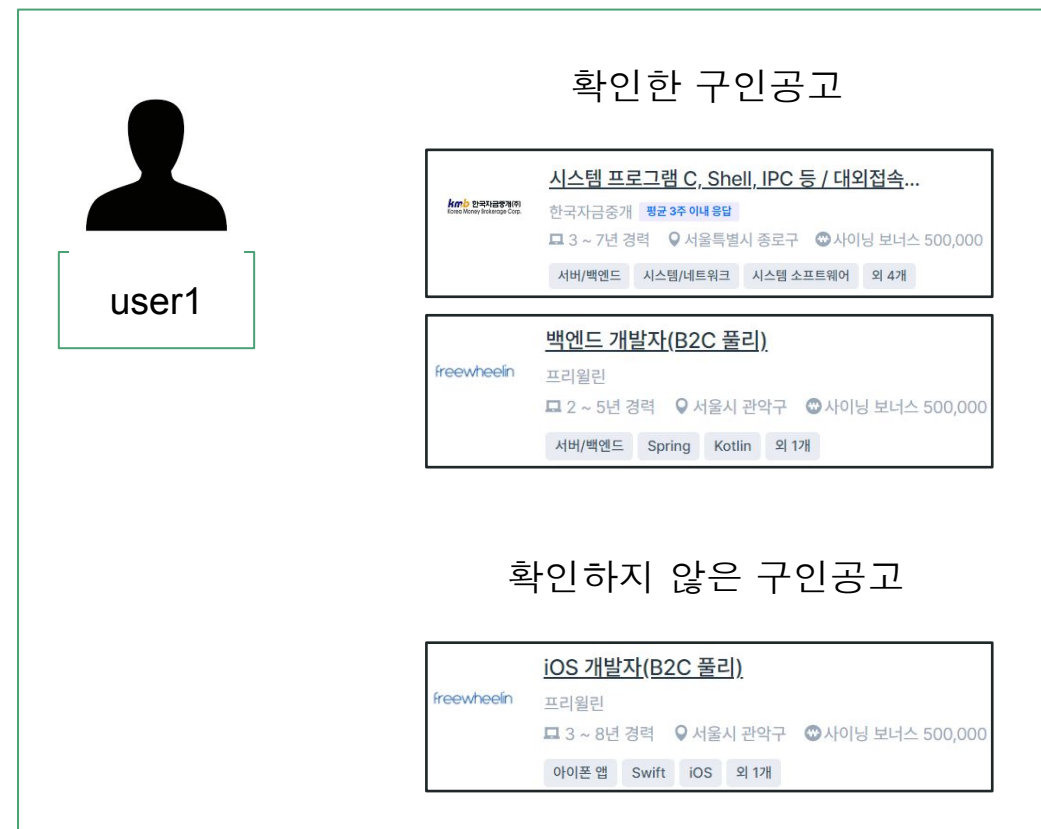
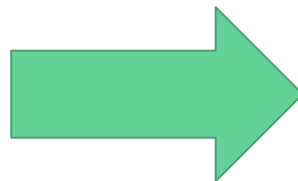
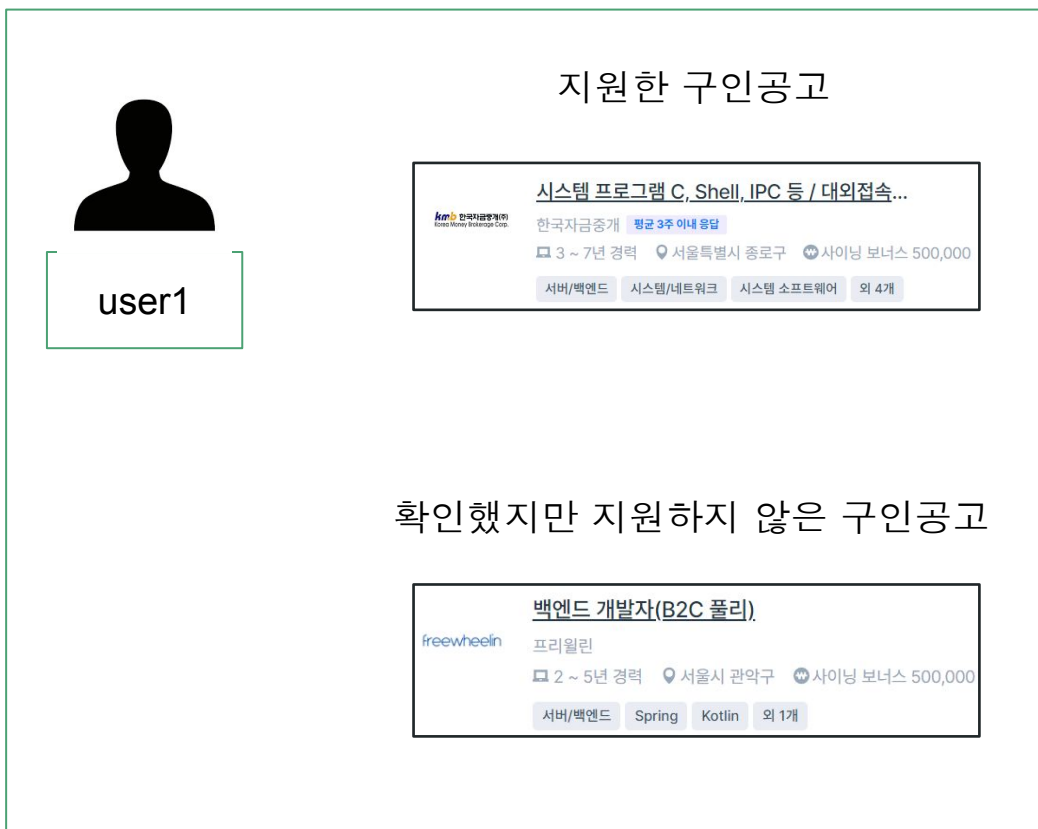
explicit data

		Items				
Users						...
	Alice	1	1	0	0	
	Bob	0	0	1	1	
	Corey	1	0	1	0	
	...					

implicit data

데이터 수정

explicit data -> implicit data 변환



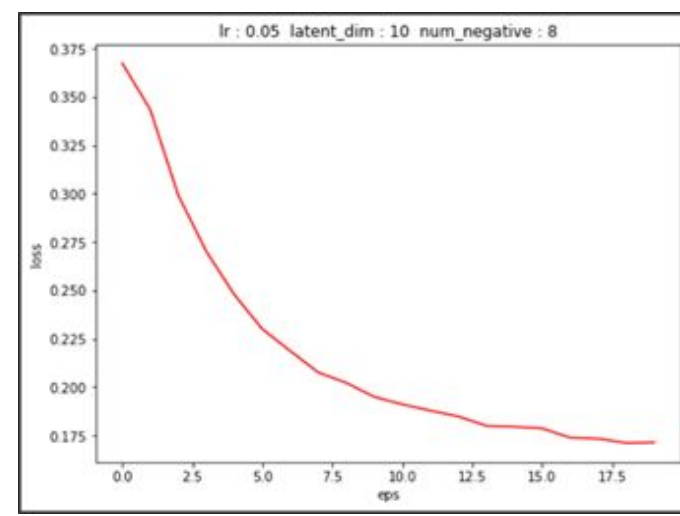
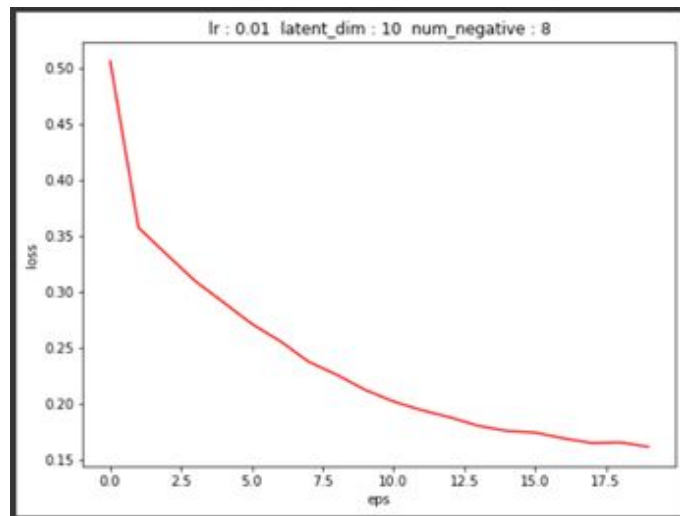
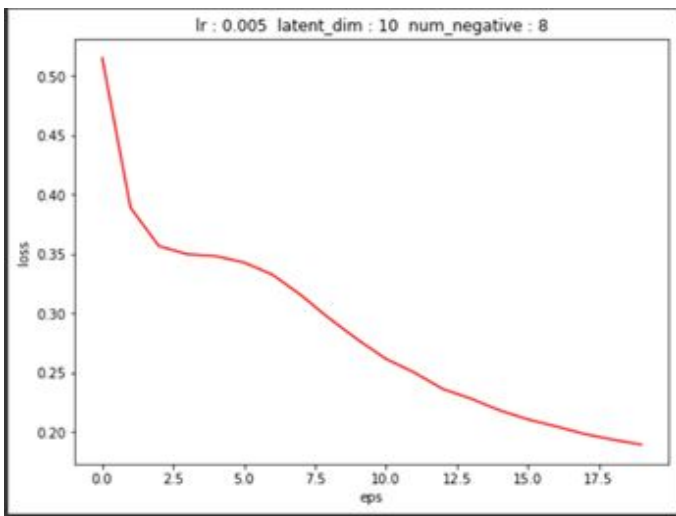
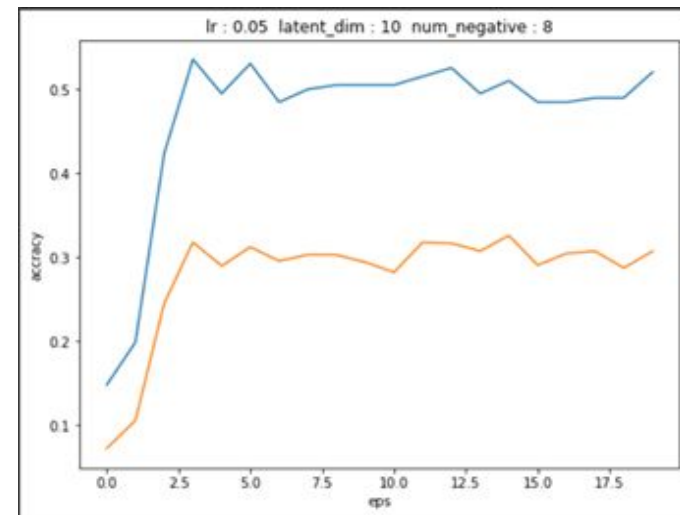
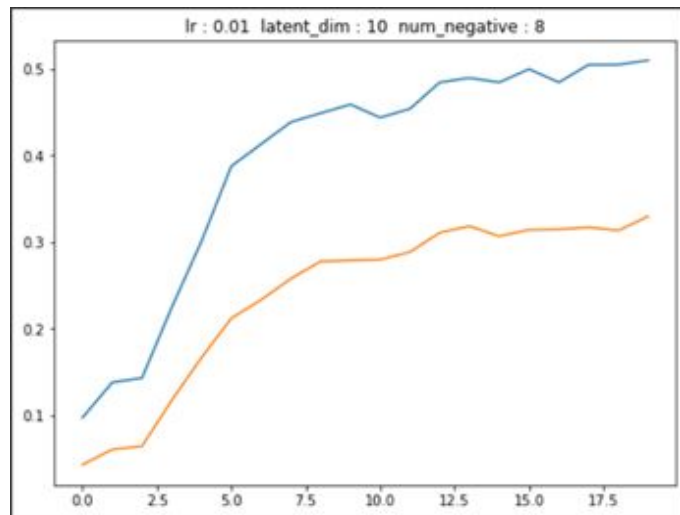
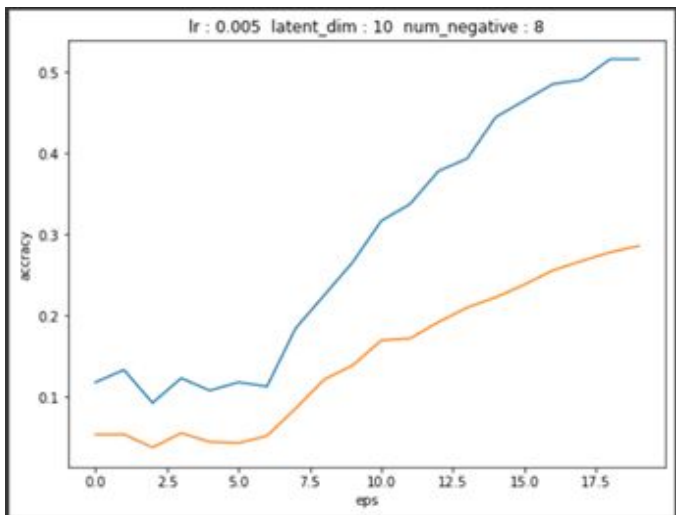
평가방법

- HR (Hit Ratio)
 - 상위 10개 항목에 대해 모델의 결과를 정답과 비교하여 일치하는 개수에 따른 점수
- NDCG (Normalizing Discounted Cumulative Gain)
 - 상위 10개 항목에 대해 모델의 결과를 정답과 비교하여 일치하는 항목의 순위에 따른 점수

loss

learn rate를 변경하면서 평가 진행

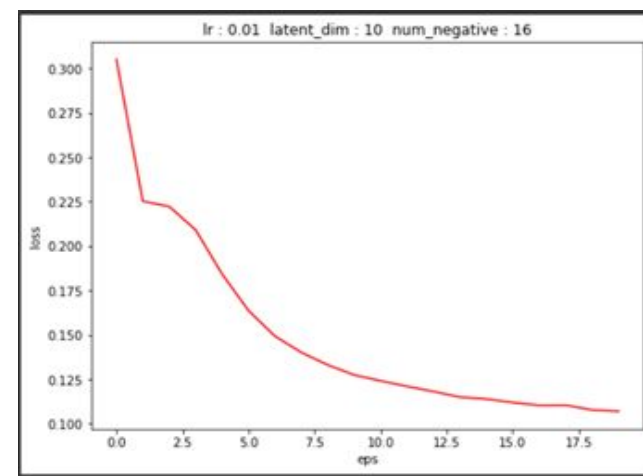
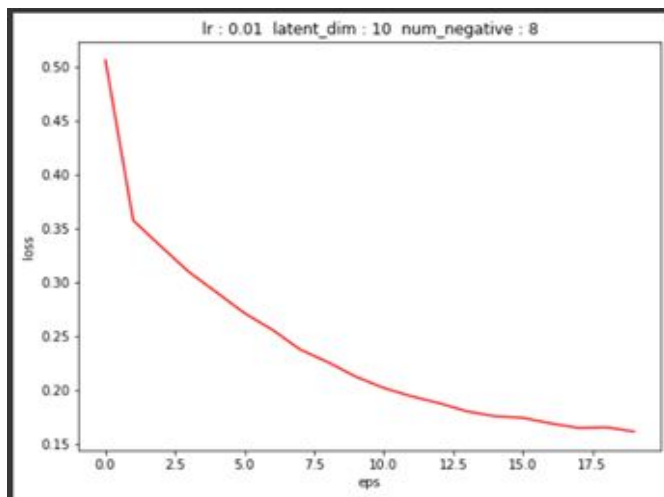
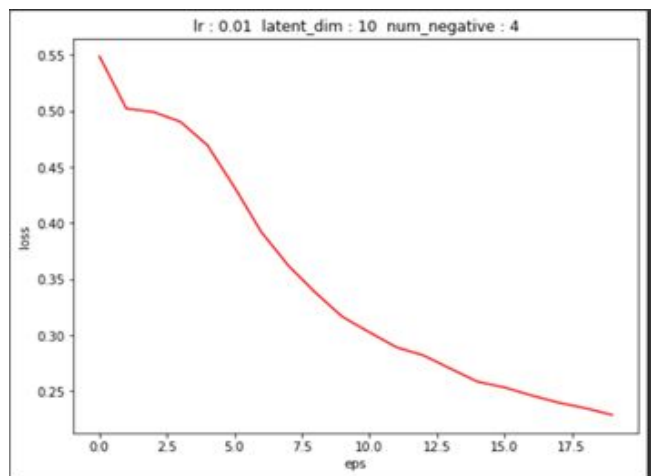
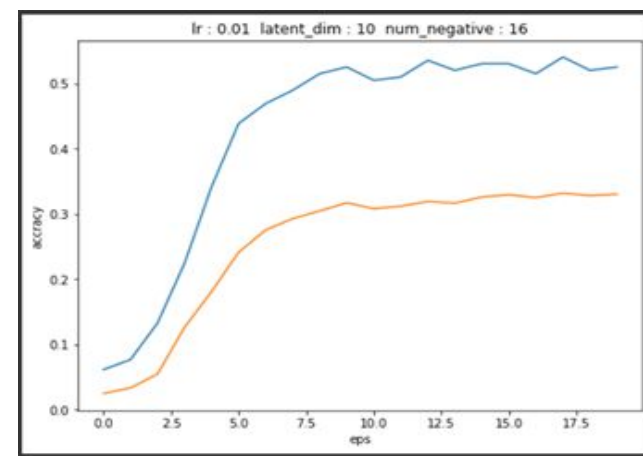
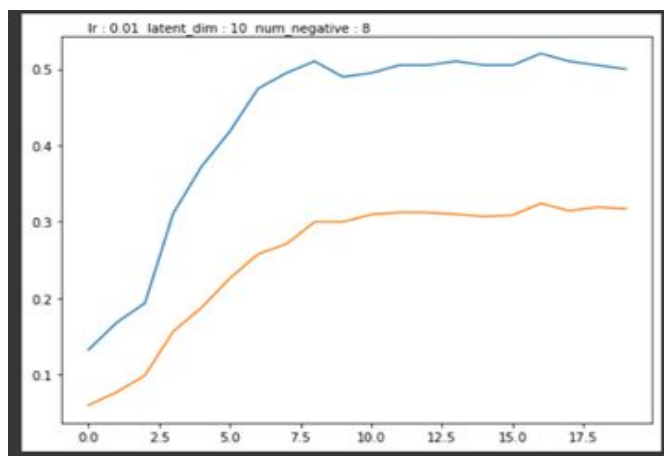
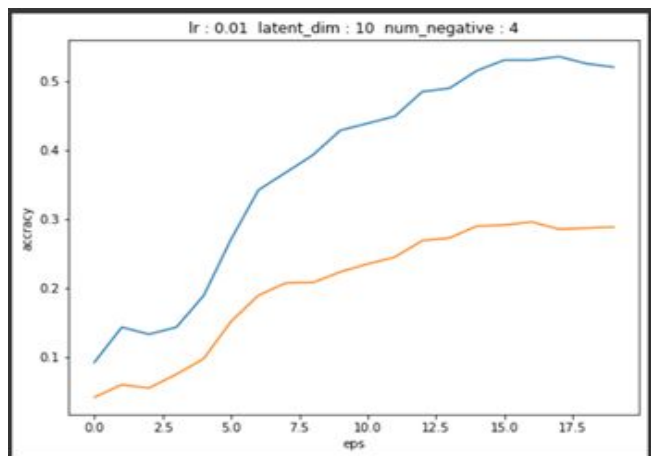
hr
ndcg



loss

negative data양을 변경하면서 평가 진행

hr
ndcg



채점 결과

STEP 4. 최종 점수로 반영할 데이터 선택하기



제출 목록	공개 점수	최종 점수 사용 데이터
ones.csv 2021-03-28 15:56:50	0.0	<input type="radio"/> 선택
output (1).csv 2021-03-28 15:56:43	81.3141683778234	<input checked="" type="radio"/> 선택
zeros.csv 2021-03-28 12:55:52	0.0	<input type="radio"/> 선택
output.csv 2021-03-28 12:47:27	79.945242984257 36	<input type="radio"/> 선택

겪었던 문제

- 과정보다는 결과 위주
 - 추천 시스템의 흐름을 일부만 이해한 채로 코드를 통해 결과만 내려고 하여 제대로 작동하지 않아 개발 시간만 늘어나게 됨.
- 데이터에 대한 부족한 이해
 - 데이터에서 각 필드가 실제로 어떤 것을 의미하는지에 대해 잘못 판단함

도움이 필요한 문제

- 프로젝트 목표에 대한 불명확함
 - 커리어 매칭이 정확히 무엇을 의미하는가
- 추가 데이터 확보
 - 질적, 양적 데이터

예정사항

- user, job에 대한 추가정보 반영(경력, 직군 등)하는 모델 적용
- 추가 확보한 데이터에서 개발자가 채용 공고를 선택하는 요인 분석

감사합니다.

부록

ncf 모델

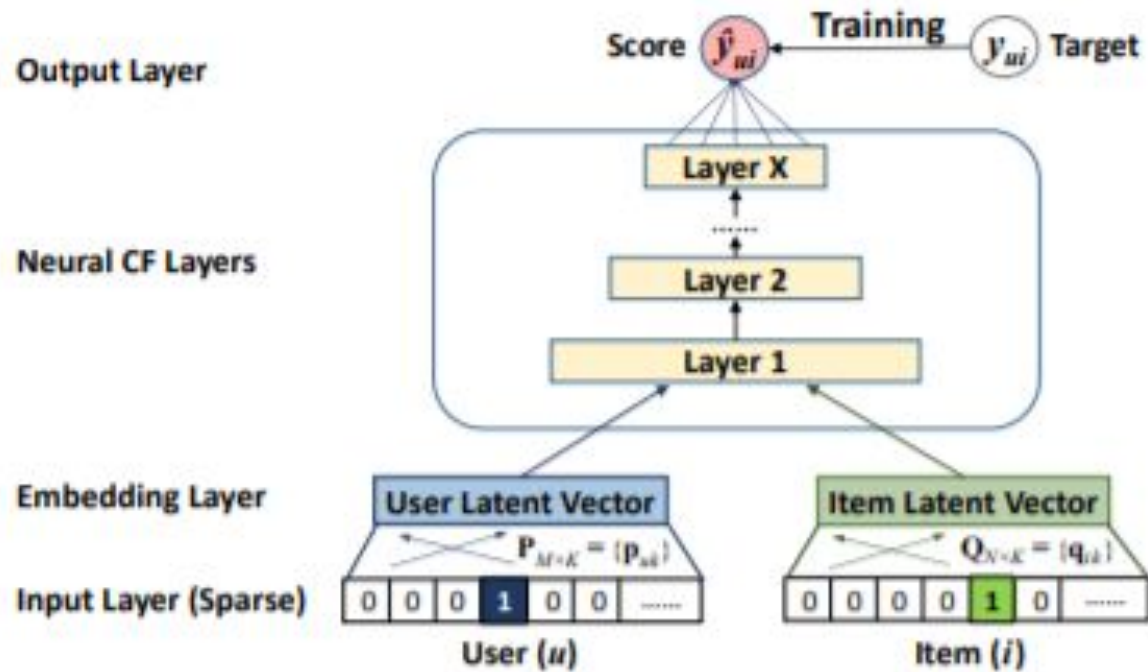


Figure 2: Neural collaborative filtering framework

평가 방법

● HR 평가방식

- 각 유저별로 test data 1개와 negative data 100개 를 이용하여 테스트
- 모델에서 나온 결과로 상위 10개 데이터에 테스트 데이터가 있을 경우 1, 아닐 경우 0으로 계산

user1 예측 데이터

job1
job2
job3
job4
...
job10

user1 test 데이터

user1, job3, 1

$$\text{HR@10} = 1$$

평가 방법

- NDCG 평가방식

- HR 방식에 test 데이터가 예측 데이터에서 몇위에 있는지에 따라 평가점수에 반영
- $ndcg = \log(2) / \log(1+x)$ x : 순위

user1 예측 데이터

job1
job2
job3
job4
...
job10

Test data

user1 test 데이터

user1, job3, 1

$$NDCG@10 = \log(2) / \log(1+3) = 0.5$$

참고자료

NCF 모델 설명 <https://arxiv.org/abs/1708.05031>