# Make-Up for You

당신을 위한 화장품 추천시스템



12기 김수아 12기 김태욱 12기 이세윤 13기 이유민

# CONTENTS

- 주제 선정 배경
  TOPIC SELECTION BACKGROUND
- 신규 유 H 에 대한 추천
  RECOMMENDATIONS FOR NEW USERS
- 1 기존 유거에 대한 추천
  RECOMMENDATIONS FOR EXISTING USERS
- 5 불과 시연
  DEMONSTRATE RESULTS
- 6 한계점 및 기대효과
  LIMITATIONS AND EXPECTATION EFFECTIVENESS

# 주제 선정 배경

TOPIC SELECTION BACKGROUD



# 문제점 1

피부타입						
전체	건성	지성				
중성	복합성	민감성				

개인이 피부를 자가 분석하는 것에 대한 어려움



# 문제점 2





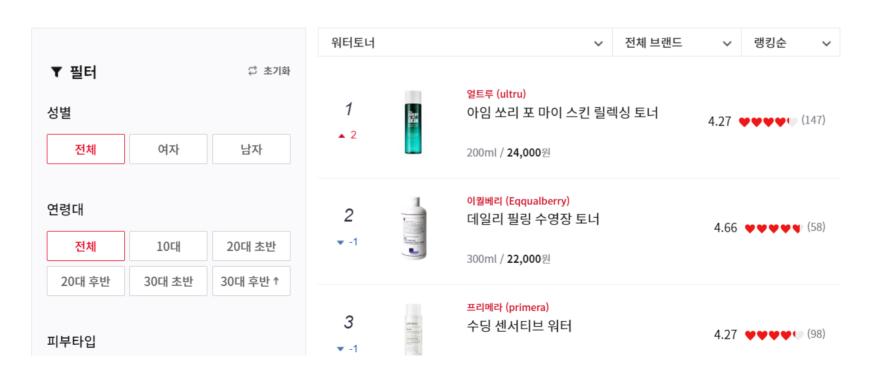


세분화된 피부타입을 고려하지 않는 현재 화장품 추천 시스템



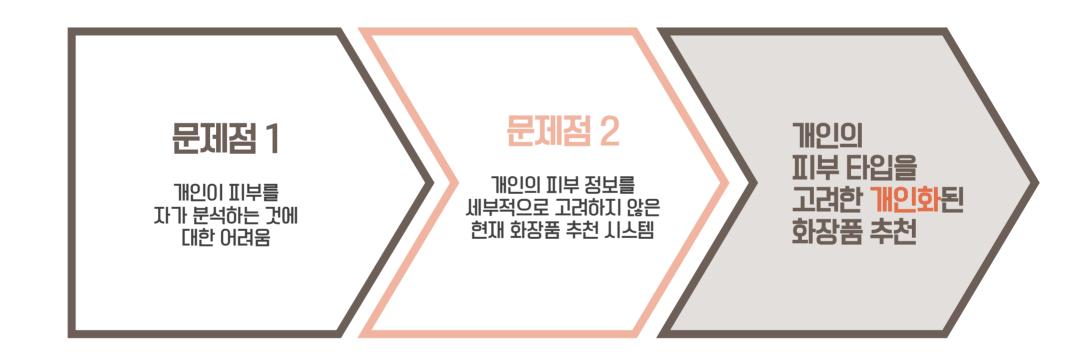
# 문제점 2

## 워터토너 랭킹



화장품 종류별 단순 랭킹만을 보여주는 사이트들





# 데이터 수집 및 전처리

DATA COLLECTION AND PREPROCESSING



진정로션 1위



## 화장품 리뷰 데이터



user\_id



23세 · 건성 · ♀

© \*\*\* Score

개인적으로 피지오겔은 뭔가 수분이 채워지는 기분이 아니라 겉도는 느낌이라서 좀 아쉬워요ㅠㅠ 건성인데도 이걸쓰면 기름진 느낌이고 이거 바르고 화장하면 화장도 겉도는 느낌이라서 좀 그렇습니다.ㅠㅠ

review

1개월 전

created\_at

피지오겔 (PHYSIOGEL)

## 카밍릴리프 AI 리페어 로션



피지오겔 (PHYSIOGEL)

브랜드홈



중간 위험도의 성분이 포함된 제품입니다.

성분정보

카테고리 로션/에멀젼 진정로션 랭킹확인

로션/에멀젼 랭킹확인

판매처 시코르, 올리브영, 랄라블라, 롭스

설명 민감 또는 과민감한 극성, 극건성 피부타입에 적합한 로션

- 건조함으로 인한 피부 자극 및 피부 스트레스를 완화시켜주고 진정시켜 피부를 건강하게 유지
- 피부의 건조함으로 인한 가려움과 자극을 완화하여 피부 장벽 강화에 도움

<u>더보기</u>

컬러/타입 200ml / 400ml

태그 보습 수분공급 저자극 피부진정

skin\_type gender





# 화장품 리뷰 데이터

	user_id	age	gender	skin_type	product_id	score	review	created_at
0	테루테루	30	female	건성	45917	5	일단 개별포장이라는 점에서 점수 후하게 주고싶음아침보단 밤에 자기전에 하 시는거 추천	2016.03.29
1	RS	23	female	건성	45917	4	이거 괜찮은데 이거 세일해서 8천원인가 해서 샀는데 싼게 아주 잘샀음 1매에 원도	2016.02.18
2	wowowo	26	female	건성	45917	3	하나씩 포장되어있어서 좋은거 같아요 얼굴에 문지르고 나서 다시 닦아야한다 는게 귀찮긴	2017.09.19
3	오늘은뭘사 지	23	female	민감성	45917	5	이거 하고나면 얼굴 부드러워짐 이거 알고난다음에 얼굴에 각질찾기 힘들어요 진짜 좋아	2017.06.18
4	져니—	29	female	중성	45917	5	시드물 녹차필링 프리메라 마일드필링 쓰고 완전 피부가 뒤집어져서 헤매다가 구매했어요	2016.12.18

글로우픽에서 938,195건의 리뷰 데이터 크롤링 진행





# 화장품 리뷰 데이터

	user_id a	ge	gender	skin_type	product_id	score	review	created_at
0	테루테루	30	female	건성	45917	5	일단 개별포장이라는 점에서 점수 후하게 주고싶음아침보단 밤에 자기전에 하 시는거 추천	2016.03.29
1	RS	23	female	건성	45917	4	이거 괜찮은데 이거 세일해서 8천원인가 해서 샀는데 싼게 아주 잘샀음 1매에 원도	2016.02.18
2	wowowo	26	female	건성	45917	3	하나씩 포장되어있어서 좋은거 같아요 얼굴에 문지르고 나서 다시 닦아야한다 는게 귀찮긴	2017.09.19
3	오늘은뭘사 지	23	female	민감성	45917	5	이거 하고나면 얼굴 부드러워짐 이거 알고난다음에 얼굴에 각질찾기 힘들어요 진짜 좋아	2017.06.18
4	져니—	29	female	중성	45917	5	시드물 녹차필링 프리메라 마일드필링 쓰고 완전 피부가 뒤집어져서 헤매다가 구매했어요	2016.12.18

유저 수 100,457명



진정로션 1위



# 화장품 정보 데이터



1개월 전



23세 · 건성 · ♀ ② 뽀



개인적으로 피지오겔은 뭔가 수분이 채워지는 기분이 아니라 겉도는 느낌이라서 좀 아쉬워요ㅠㅠ 건 성인데도 이걸쓰면 기름진 느낌이고 이거 바르고 화장하면 화장도 겉도는 느낌이라서 좀 그렇습니 С. тт

피지오겔 (PHYSIOGEL)

#### product\_nm

### 카밍릴리프 AI 리페어 로션



200ml / **0**원

amount price

average\_score

피지오겔 (PHYSIOGEL) **brand** 

브랜드홈

성분정보

4.27



판매처

중간 위험도의 성분이 포함된 제품입니다.

ingredients

로션/에멀젼 진정로션 카테고리

category\_detail category

로션/에멀젼

시코르, 올리브영, 랄라블라, 롭스

민감 또는 과민감한 극성, 극건성 피부타입에 적합한 로션

- 건조함으로 인한 피부 자극 및 피부 스트레스를 완화시켜주고 진정시켜 피부를 건강하게 유지
- 피부의 건조함으로 인한 가려움과 자극을 완화하여 피부 장벽 강화에 도움

<u>더보기</u>

컬러/타입 200ml / 400ml

태그

수분공급

저자극

tags





## 화장품 정보 데이터

제품 ID, 제품명, 브랜드, 카테고리, 세부카테고리, 가격, 용량, 성분, 위험도 등의 정보를 크롤링

> t json II일

```
{
  "product_id": 45917,
  "title": "[단종] 청보리 필링 패드",
  "brand": "이니스프리 (innisfree)",
  "category": "각질케어",
  "category_detail": "각질케어 필링패드",
  "price": "12,000원",
  "amount": "20p",
  "ingredients": [
  {
  "title": "정제수",
  "purpose": [
  "II부컨디서님제,용제"
],
  "hazard_class": "green",
  "hazard_class_num": 1
},
```





## 화장품 정보 데이터

제품 ID, 제품명, 브랜드, 카테고리, 세부카테고리, 가격, 용량, 성분, 위험도 등의 정보를 크롤링

json II일

### 제품 수 33,314개

```
"product_id": 45917,

"title": "단종J 정모리 필당 패드",

"brand": "이LI스프리 (innisfree)",

"category": "각질케어",

"category_detail": "각질케어 필링패드",

"price": "12,000원",

"amount": "20p",

"ingredients": [
{

"title": "점제수",

"purpose": [
"II부컨디너님에, 용제"
],

"hazard_class": "green",

"hazard_class_num": 1
},
```



# 기본 전처리

- 나이: 80세 이상 / 10세 이하인 데이터 제개
- ID 결측치 제귀
- 제품 성분 정보 없는 데이터 제가

## 필요없는 데이터 제거

- 필요한 화장품 카테고리만 선택

## 최종 카테고리

- 각질궤어
- 메이크업클렌저
- 페이셜클렌저
- 에센스/세럼
- 스킨/토너
- 파운데이션
- 크림
- 시트마스크
- 선크림
- 바디오일/밤
- 립궤어
- CC크림
- 베이스/프라이머
- 바디로션/크림 남성 클렌징
- 페이스마스크

- 로션/에멀젼
- IILI시I마우더
- 미스트
- 선스프레이
- 배쓰/샤워
- 메이크업픽서
- 에프터선
- BB크림
- 바디미스트
- 베이비 선궤어
- 선쿠션
- 선베이스

  - 베이비 클렌징



## 리뷰 데이터 전처리

- 정규표현식으로 특수문자, 영어, 숫자 삭제

# 단어 사전 생성 및 Tokenizing

- 브랜드명, 상품명 등의 고유 명사를 잘 추출해 내기 위한 사용자 단어 사전 생성
- Mecab을 이용한 토크나이징
- \* 원하는 단어에 대해 단어 비용을 낮춰 사용자 단어사전을 우선적으로 토크나이징에 적용

시드물녹차필링,프리메라마일드필링쓰고 완전피부가뒤집어져서헤매다가구매했어...

['시드물', '녹차', '필링', ',', '프리메라', '마일드', '필링', '쓰', '고', '완전', '피부', '가', '뒤집', '어', '져서', '헤매', '다가', '구매', '했', '어', '.', '..']





# 최종 데이터 632,443건

	user_id	age	gender	skin_type	product_id	score	review	created_at	review_token
0	테루테루	30	female	건성	45917	5	일단 개별포장이라는 점에서 점수 후하게 주고싶음아침보단 밤에 자기전에 하 시는거 추천	2016.03.29	['일단/MAG', '개별/NNG', '포장/NNG', '이/VCP', '라는/ET
1	RS	23	female	건성	45917	4	이거 괜찮은데 이거 세일해서 8천원인가 해서 샀는데 싼게 아주 잘샀음 1매에 원도	2016.02.18	['이거/NP', '괜찮/VA', '은데/EC', '이거/NP', '세 일/NNG',
2	wowowo	26	female	건성	45917	3	하나씩 포장되어있어서 좋은거 같아요 얼굴에 문지르고 나서 다시 닦아야한다 는게 귀찮긴	2017.09.19	['하나/NR', '씩/XSN', '포장/NNG', '되/XSV', '어/EC',
3	오늘은뭘사 지	23	female	민감성	45917	5	이거 하고나면 얼굴 부드러워짐 이거 알고난다음에 얼굴에 각질찾기 힘들어요 진짜 좋아	2017.06.18	['이거/NP', '하/W', '고/EC', '나/WX', '면/EC', '얼 굴/
4	져니—	29	female	중성	45917	5	시드물 녹차필링 프리메라 마일드필링 쓰고 완전 피부가 뒤집어져서 헤매다가 구매했어요	2016.12.18	['시드/NNG', '물/NNG', '녹차/NNG', '필링/NNG', '프리/NN

유저 수 89,425명

제품 수 17,185개



# 피부 타입 별 분석







건성

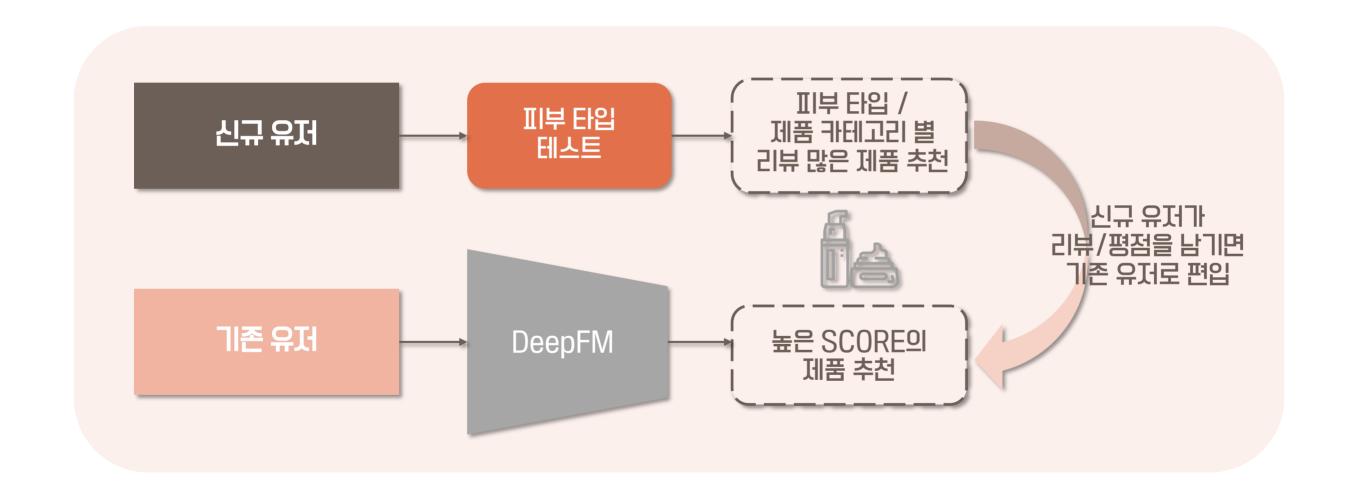
IJ성

민감성

# 추천시스템 구조

• RECOMMENDATION SYSTEM STRUCTURE





# 신규 유저에 대한 추천

RECOMMENDATIONS FOR NEW USERS



## 준비 과정

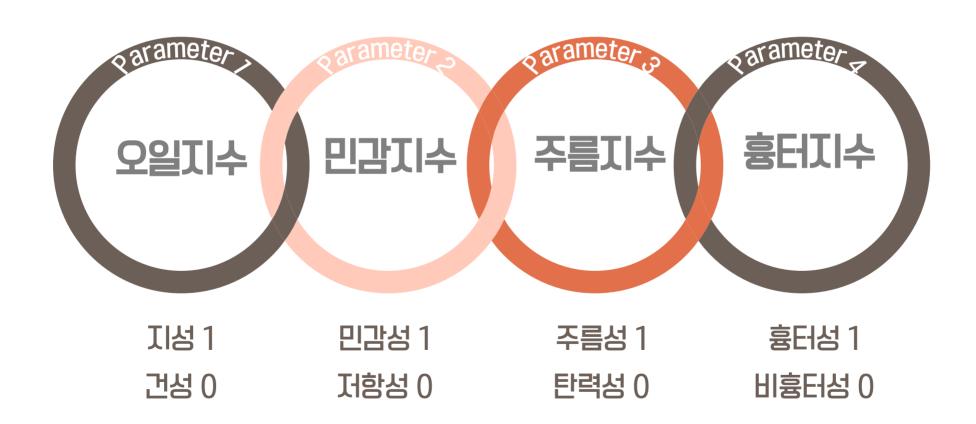
기존 유저와 달리 신규 유저는 구매한 상품 정보 및 평점 정보가 없다

이러한 신규 유저에 대한 cold start problem을 해결하기 위해 사이트 회원 가입 시 <mark>피부 타입 테스트</mark> 진행

피부 타입 테스트 기반 분류

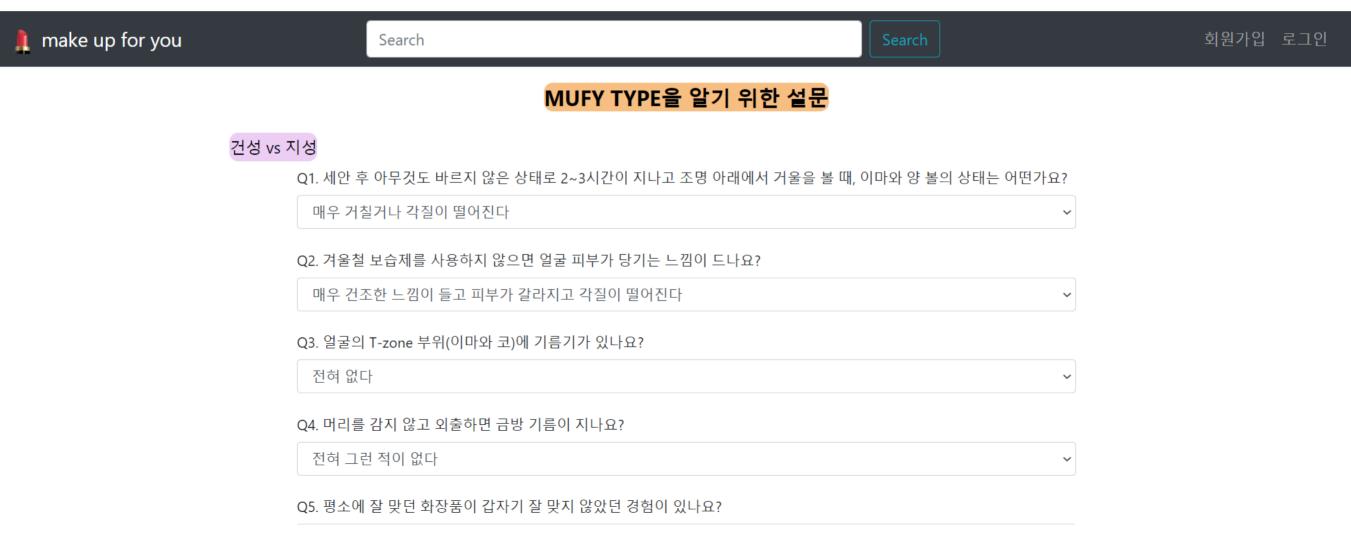


# 4가지 파라미터를 기준으로 16가지 피부 타입 카테고리 생성



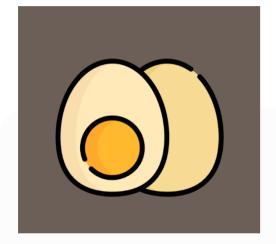


# 총 19가지 질문을 통해 16가지 피부 타입 중 하나로 유저 분류

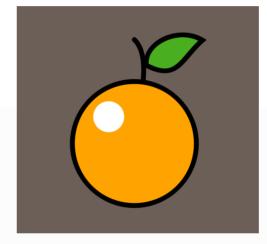




## IJ성









까달걀

#건강 #지성

제주감귤

한라봉

큰 문제가 되는 트리블이 없고 피부톤도 일정한 지성 II부

#색소침착 #지성

메추리알

햇빛에 쉽게 타거나, 잡EI와 같은 색소침착이 걱정되는 III부.

EI존은 기름기가 넘치는데 비해 수분은 부족하지만 매우 민감하지는 않아 다행

#### #주름 #지성

피부는 건강하지만 탄력이 없는 지성 피부. 민감하지는 않기 때문에 트러블과 흉터의 위험으로부터는 비교적 안전한 Ⅲ부

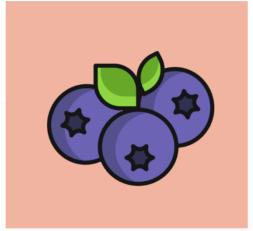
#### #색소침착 #주름 #지성

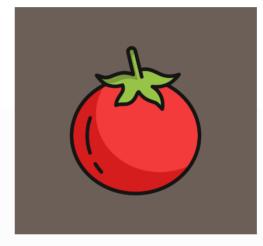
피부는 건강하지만 탄력이 없고 트리블 흉터가 남아 있는 피부. 민감하지는 않기 때문에 탄력관리와 흉터관리가 중요!

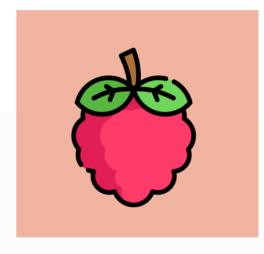


## 민감 지성









석류

블루베리

찰토마토

산딸기

#### #민감 #지성

번들거림과 여드름이 공존하는 지성피부지만 색소침착의 위험에서는 다소 벗어나 있음

#### #색소침착 #민감 #지성

여러 피부 트러블이 반복해서 일어나는 민감한 지성 피부. 다양한 트러블이 일어나는 악순환을 끊을 대책이 시급!

#### #주름 #민감 #지성

탄력이 없어 비교적 모공이 늘어져 보이지만 아직 여드름 흉터와 같은 색소침착이 눈에 띄게 진행되지는 않은 피부

#### #색소침착 #주름 #민감 #지성

색소침착이 이미 진행된 데다 탄력까지 떨어진 민감한 지성 피부. 자주 발생하는 트러블로 인해 피부가 이미 지쳐 있는 상태



# 건성









애기피부

말광량이 뻬뻬

땅콩

멜론

#### #건강 #건성

특별히 신경쓸 게 없는 건강한 피부로, 피부트러블 예방을 위해 보습 정도만 미리 신경쓰기

#### #색소침착 #건성

기미와 주근배 등의 잡티를 가지고 있지만 민감도가 높지 않고 모공이 보이지 않는 비교적 건강한 피부

#### #주름 #건성

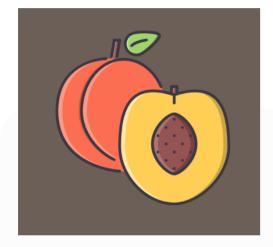
탄력이 약간 떨어지지만 건강한 건성 피부. 민감하지 않으니 트러블로 인한 흉터의 위험으로부터 는 비교적 안전!

#### #색소침착 #주름 #건성

피부상태는 비교적 건강하지만 트리블 흉터가 남아 있는 건성피부. 다행히 환경변화에 민감하지는 않으니 흉터관리에 신경써주자



## 민감 건성









복숭아

수박

파인애플

낙엽

#### #민감 #트리블 #건성

수분이 부족해서 발생하는 여드름의 위험이 높지만, 발생한 트러블은 컨디션만 좋다면 잘 완화되는 어린 피부

#### #민감 #색소침착 #건성

색소침착과 함께 노화가 이미 진행된 피부일 확률이 높음. 유수분이 부족하여 나이가 들수록 점점 민감해질 수 있으나 지금 당장 민감하지는 않은 피부

#### #주름 #민감 #건성

주름이 있는 얇고 노화된 피부. 건조하고 탄력이 부족하여 주름이 두드러져 보일 수 있으니 주의가 필요

#### #주름 #민감 #색소침착 #건성

피부가 얇아 민감하고 색소침착과 주름의 위험도 있어 종합적인 대책이 필요한 피부



# 자신의 II부 타입, 관심 있는 카테고리를 기반으로 리뷰가 가장 많은 10개의 제품 추천

및 make up for you Search Search

크림▼

메추리알 타입을 위한 추천 ▼

검색

1위



#### [단종] 레드 이레이징 크림

메디큐브 (medicube)

크림

주근깨 등 과색소침착에 대한 임상 시험을 완료한 민감성 피부에도 사용할 수 있는 저자극 피부 장벽 강화 크림입니다.-주근 깨 등 피부 색소 침착 개선-피부 수분 손실량이 개선되어 피부 보습막 형성-저자극 제품으로 민감성 피부도 사용 가능-뛰어 난 보습력과 피부 각질 개선-피부 미백 및 주름 개선에 대한 기능성 보고 완료

2위



#### 화이트홀릭

더블유랩 (W.Lab)

크림

나 원래 하얘! 타고난 척, 쌩얼인 척, 원래 하얀 척. 바르는 즉시 도화살 피부로 만들어주는 1초 피부톤 튜닝 크림- 즉각적인 화 이트닝- 미백, 탄력- 잡티 완화- 피부 보습



## 신규 유저가 리뷰와 평점을 남겨 기존 유저로 편입

#### 백설기 크림



#### 김태욱 메추리알 3점

무난했어요. 그런데 그닥 효과가 확 나타나지는 않아서 누가 선물해준다면 다시 쓸 듯?

### 투엑스 퍼스트 에센스



#### 김태욱 메추리알 4점

돈만 많으면 계속 쓸 제품... 향도 좋고 촉촉하고 기초 첫 단계로 바르기 좋아요 완전 묽지는 않아서 스킨 대용으로 쓰면 좋을 것 같아요

### 레드 이레이징 크림



#### 김태욱 메추리알 5점

이거 만드신 분 들숨에 건강을 날숨에 재력을 얻으시길 바랍니다. 저녁에 듬뿍 발라주고 자면 다음날 아침 피부 광명을 찾을 수 있어요.

### 센텔라카 스팟



#### 김태욱 메추리알 4점

여드름 날 때마다 이거 써요 소량만 발라줘도 진정시켜줘서 좋네요 근데 흉터 자체가 없어지지는 않아서 1점 뺐어요

# 기존 유더에 대한 추천

RECOMMENDATIONS FOR EXISTING USERS



## 준비 과정

기존 유저 귀워드 기반 피부 타입 분류 진행

- 토크나이징 및 EDA를 통해 얻은 리뷰 키워드를 통해 룰베이스 방식으로 기존 유저의 피부 타입 분류

II부 타입	유저 수
애기III부	14,596
블루베리	12,385
석류	11,479
깐달걀	10,899
복숨아	9,735
산딸기	8,320
수박	5,191
메추리알	5,064
말광량이뻬뻬	4,480
낙엽	2,750
찰토마토	1,512
한라봉	754
파인애플	735
땅콩	528
제주감귤	503
멜론	494

총 89,425명



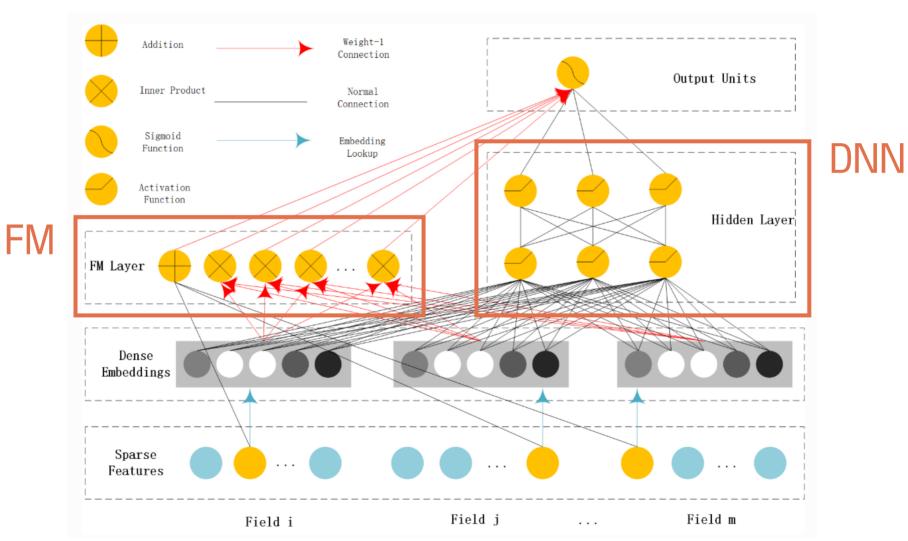
# 준비 과정

모델에 구성 성분 및 태그 넣기 위해 PCA를 통한 데이터 차원 축소



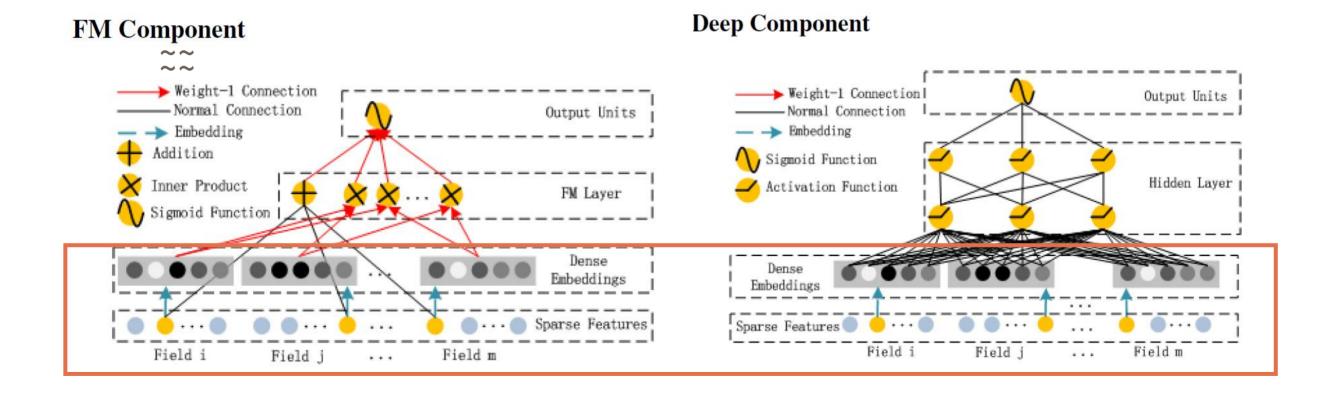


# DeepFM = FM + DeepLearning



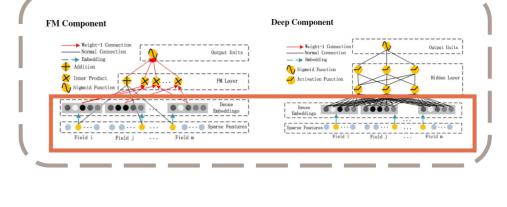


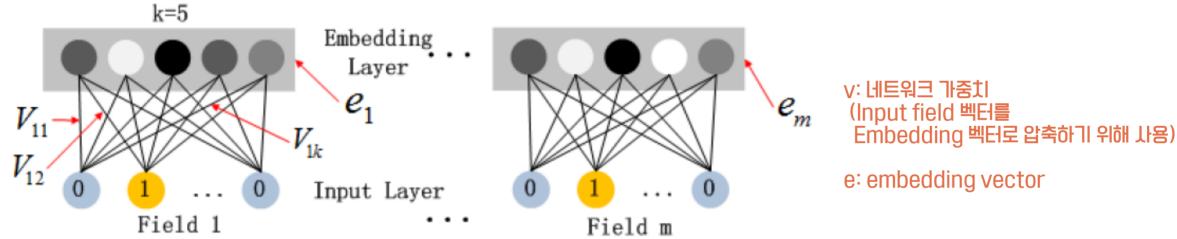
# DeepFM





# **Embedding Layer**





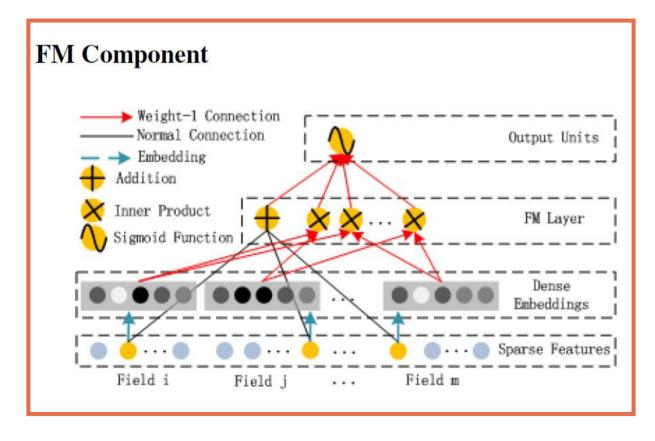
One-Hot 벡터로 변환하여 sparse feature로 나옴 (field)

총 m개의 변수 (m=536)

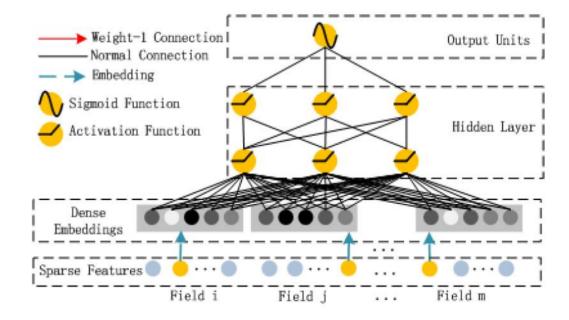
서로 다른 input 변수 벡터의 길이가 다르지만, k개의 차원의 같은 크기로 임베딩 (k=4)



## DeepFM



### **Deep Component**



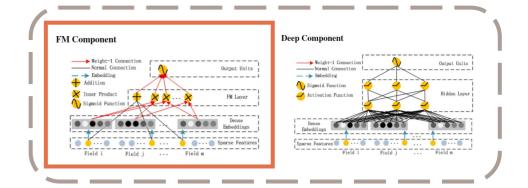


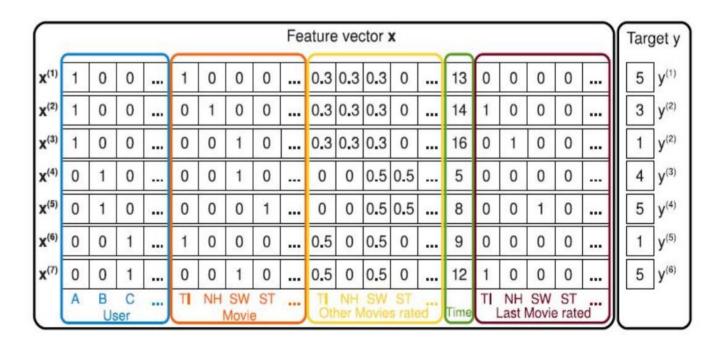
## Factorization Machines (FM)

한 햄에는 하나의 User, 하나의 Item

FM 모델은 각 상호작용에 대해 파라미터를 그대로 사용하는 것이 아니라, 이를 분해하여 사용

\* 희소한 데이터임에도 불구하고 고차원의 상호작용에 대한 훌륭한 Ih라미터 추정치를 산출할 수 있는 중요한 역할을 함

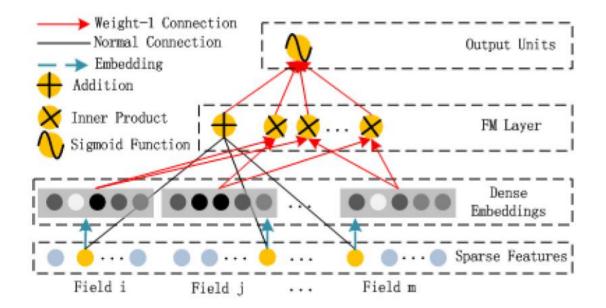


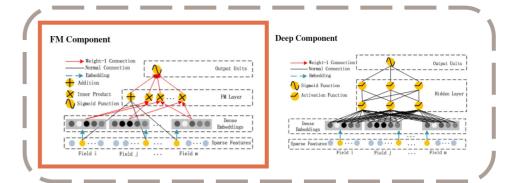




## Factorization Machines (FM)

### **FM Component**





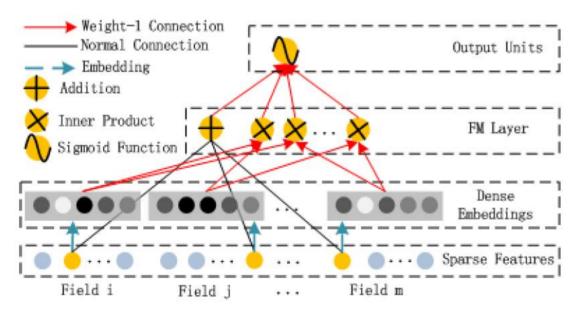
Latent vector Vi : 다른 피겨들과의 interaction의 영향을 측정

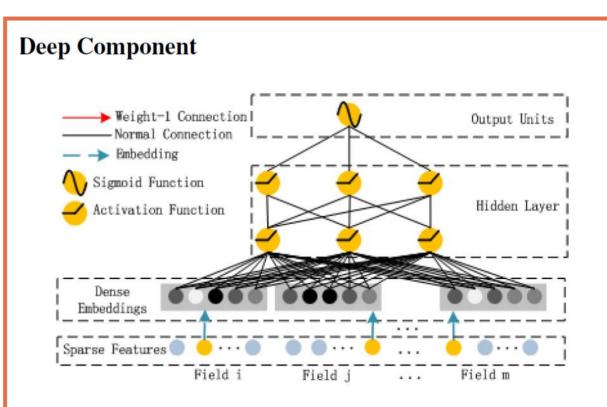
FM에서는 저차원의 상호작용을 측정 (2차원)



## Deep Component

## **FM Component**

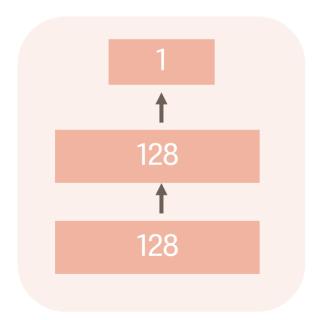


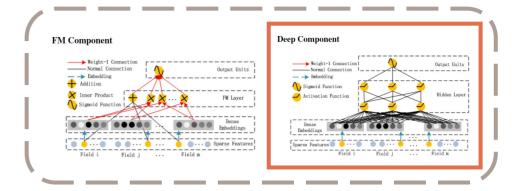




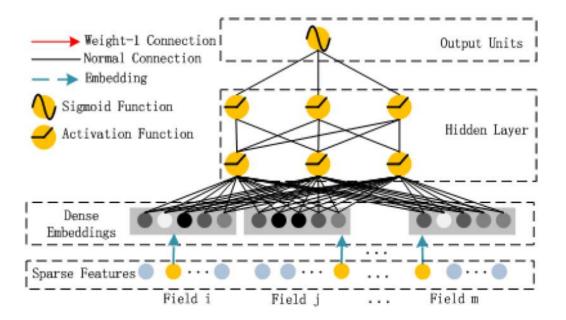
## Deep Component

Deep에서는 고차원의 상호작용 측정





### **Deep Component**





FM: 복잡도가 높아 보통 feature 2개의 상호작용만 고려

Wide&Deep: Wide (linear) + Deep (NN) = Wide Part에 Feature Engineering 필요

## DeepFM은

고차원, 저차원의 상호작용을 모두 고려하기 때문에 알기 어려운 implicit한 상호작용을 찾아낼 수 있으며, feature engineering이 필요 없음



# Output

	user_id	product_id	product_nm	pred_score
0	테루테루	7618	[단종]예현 진본 수면 크림	5.213403
1	테루테루	8458	메리고라운드 리터닝 나이트크림	5.163582
2	테루테루	9950	멀티 액티브 나이트 유쓰 리커버리 크림	5.163582
3	테루테루	14673	피오니 브라이트닝 안티에이징 크림	5.147834
4	테루테루	4551	타임와이즈 플러스+ 레지나-펌™ 나이트 트리트먼트	5.144683
5	테루테루	12832	미네랄 레디언스 오버나이트 디 스트레싱 크림	5.140707
6	테루테루	15184	에너지 부스트 나이트 리커버리밤	5.138083
7	테루테루	14419	[단종]끄렘므 레빠라뜨리스 리스토렉티브 페이셜 크림	5.133038
8	테루테루	3875	플레지엄 미드나이트 스페셜 크림	5.129218
9	테루테루	14890	비타 A 크림 포르테	5.129203
10	테루테루	15854	이데베논 모이스춰라이징크림	5.127507
11	테루테루	10094	씨에이에프 리페어 크림	5.125990
12	테루테루	3431	압솔뤼 프리미엄 베타엑스 나이트 크림	5.125675
13	테루테루	14131	레드 진생 수면 크림	5.124702
14	테루테루	10114	윤로 수결 탄력크림	5.124627
15	테루테루	6438	[단종]비윤진 고진 순금 크림	5.123224

nake up for you

Search

Searc

로그아

크림▼

당신을 위한 추천 ▼

검색

1위



#### 라스트 솔루션 제트 크림

#### 인쏘뷰 (insobeau)

크림

수분공급, 안티에이징, 주름, 미백, 탄력, 피부정돈, 모공관리의 7가지 기능을 하나에 담은 크림. 피부에 쏙 스며드는 명품텍스 처인 리포좀 공법이 업그레이드 되어 유효성분들이 더욱 효과적으로 피부에 침투합니다.

2위



#### 수퍼 레스토러티브 토탈 아이 컨센트레이트

#### 클라랑스 (CLARINS)

크림

어려보이는 눈가를 위한 아이 컨센트레이트,식물성 농축액과 유효 성분이 눈가 전체를 건강하게 가꾸어 줍니다.아이 컨센트 레이트로, 세월을 거스르는 듯한 아름다운 눈가를 선사합니다.

39



#### 모이스트3 (큐브) 크림

#### 입큰 (IPKN)

크림

천연유래 아미노산 인증 오일이 한번 더 감싸줘 당김ZERO- 모이스트 라인은 독자적인 리퀴드 아이스와 특허받은 비타민 리 포솜이 부여하는 부드럽고 촉촉한 흡수력과 달맞이꽃, 아르간에서 추출한 에코서트 내추럴 오일 함유로 탄력있는 속보습 케



• DEMONSTRATE RESULTS

# 한계점 및 기대효과

LIMITATIONS AND EXPECTATION EFFECTIVENESS



## 기대효과

- 1. 설문을 통해 비교적 간단하고 직관적인 피부타입 정보를 사용자에게 제시
- 2. 개인의 피부 상태와 화장품 구성성분 등을 고려한 개인화된 화장품 추천시스템 구축

## 한계점

- 1. 부정적 반응에 대한 데이터 부족
- 추천 순위가 낮은 제품과 쓰면 안되는 제품은 다르지만 따로 구분하지 못함
- 2. 즉각적 정보 반영 불가
- 3. 룰베이스 기반 유저 피부 타입 분류로 낮은 확장성

## 개선방향

- 1. 화장품 성분 간의 상호작용을 고려한 화장품 라인 추천
- 2. 쓰면 안되는 제품에 대한 정보 추가 제공

# Reference

## 도움주신 분

11기 김대웅

## 리뷰 데이터

글로우픽 https://www.glowpick.com/

## DeepFM

DeepFM: A Factorization-Machine based Neural Network for CTR Prediction

https://arxiv.org/abs/1703.04247

https://greeksharifa.github.io/machine\_learning/2020/04/07/DeepFM/

https://dos-tacos.github.io/paper%20review/DeepFM/

## 피부타입

16drops http://16drops.com/

바우만 II부타입 https://skintypesolutions.com/

# Q&A

• QUESTION AND ANSWER





12기 **김수아** # 산딸기

"낭만개발 김사부"



12기 김태욱

# 메추리알

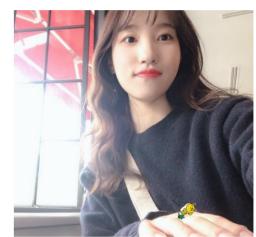
"오늘도 뜨는 해를 본다..."



12기 이세윤

# 애기피부

"데이터사이언티스트를 꿈꾸는"



13기 이유민

# 석류

"작고 소중한 우리 팀 ♥ "

# 감사합니다

