



# Perfume Recommendation System



시계열자료분석팀    오정민 강현주 김민정 배지현



# Contents

---

향수 추천시스템

01

02

03

04

05

지난주 복습

데이터 수정

유사도

추천시스템

한계/의의



지난주

데이터 수정

유사도

추천시스템

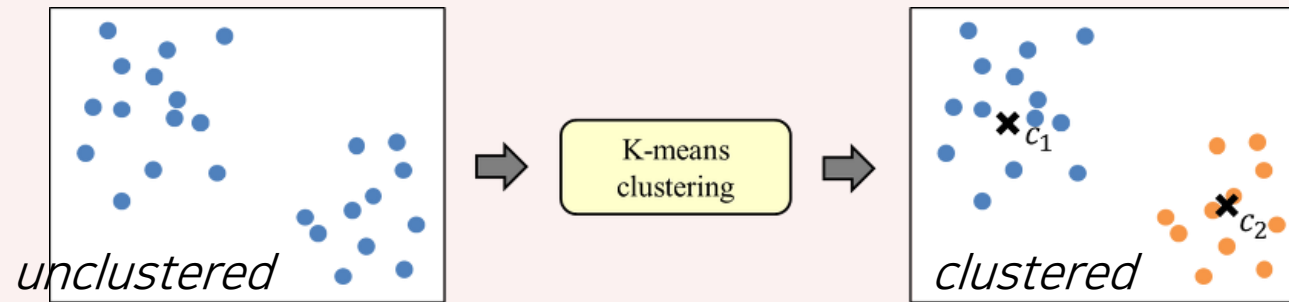
한계/의의

## 03. 유사도

색   특징   향   최종



### K-means clustering



이미지 영역을 동질의 시각적 영역 단위로 나누는 작업

- 주어진 데이터를 여러 그룹으로 나누는 방법
- 각 그룹의 중심과 그룹 내의 데이터와의 거리의 제곱합을 최소화하는 방향
- 각 데이터 오브젝트의 소속 그룹을 업데이트 해 줌으로써 클러스터링 수행



지난주

데이터 수정

유사도

추천시스템

한계/의의

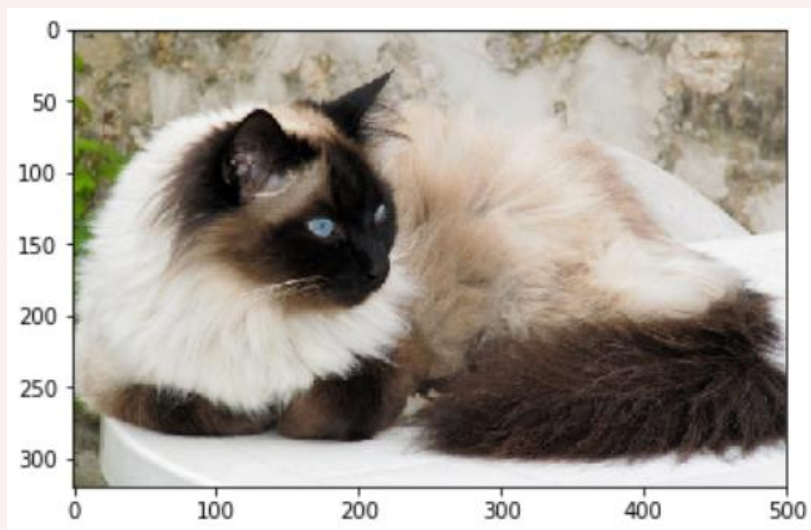
## 03. 유사도

색 특징 향 최종



### K-means clustering

이미지 영역을 동질의 시각적 영역 단위로 나누는 작업





지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색   특징   향   최종



색 유사도

R	G	B
55	132	100
24	76	215
...	...	...



최종적으로 구한 1가지 대표색을 RGB 값으로 변환한 후  
유클리디안 거리를 통해 유사도 계산!



## 03. 유사도

색 특징 향 최종



왜 유클리디안 유사도를 썼나?

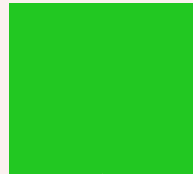
지난주

데이터 수정

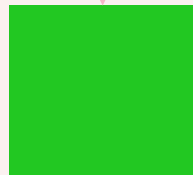
유사도

추천시스템

한계/의의



↑  
same  
color  
↓



코사인	0.783
유클리디안	191

코사인	0.783
유클리디안	128

서로 다른  
색인데도  
같다

코사인보다 유클리디안이 색깔의 차이를 더 잘 나타낸다!



지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색 특징 향 최종



최종 특징 데이터

Name	Sex	Cold	Hot	fall	Spring	Day	Night	Long	Sil	wrate
Sweet Tobacco Spirits	M	100	4.84	91.51	24.24	58.78	92.12	4.04	2.68	4.23
...	...	...	...	...	...	...	...	...	...	...



각 열별로 scaling 해준 후 **유클리디안 거리**를 통해 계산!





지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색 특징 향 최종



### Main Accords & Notes

노트 간의 관계를 고려해서 공간 벡터에 향 위치시키기

# HOW?

## 리뷰 데이터를 사용

리뷰에서 묘사된 **향의 문맥**을 통해  
노트 간의 관계를 고려해보자!

From beginning to end it is a blossoming cold jasmine flower, accompanied with a side of warm, oozing, radiant amber and served on a smooth, creamy wooden plate. The lasting power and sillage is amazing! 10/10 on skin and a 12/10 on clothing where it will last for weeks, if not months! There is no need for more than 1-2 sprays so it can last you a very long time (unless you're like me and spray it constantly on yourself just because you can lol). I also very much enjoy the fact that there are refillable bottles and therefore it is economical and eco-friendly. This is my MOST complimented perfume to date with both men and women of all ages, but in my case, especially the men!

사람들은 리뷰를 쓸 때  
본인들이 향에 관해  
느낀 것들을 묘사한다





지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색 특징 향 최종



Review Data

Word2Vec (with 긍·부정 리뷰 11,315개)

name	pos_review	neg_review
Amber Oud Gold Edition	[review al h oud gold edit smell like pla	결과값이... 기준1. 앞서 나왔던 '향에 대한 설명'과 부합하는지 기준2. Fragrance Wheel과 부합하는지
First Instinct	['sleeper hi stars fantas luxurious bottle and covere	disappointed! bought this...
Ambra	['i struggle to understand the exactly identical to colonia	in scent and performance it is just a new...

이 중 임베딩 결과가  
가장 좋은 모델을 채택하자

Word2Vec 모델 학습

차원을 달리하여 네 개의 모델을 만듦(20/30/40/50차원)

window = 10, min\_count=30, workers=4, iter=100, sg=1 sample = 1e-3

Skip-Gram 사용



지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색 특징 향 최종



Word2Vec 모델 평가 예시

“비슷한 성질을 지닌 향은 가까이 위치해 있을 것이다”

green

herbal, floral, woody,  
spicy, moss 등

Nearest points in the original space:

moss	0.177
herbal	0.247
citrus	0.305
woody	0.339
floral	0.351

Nearest points in the original space:

herbal	0.261
moss	0.295
floral	0.354
citrus	0.380
musky	0.424

20차원

30차원

40차원

50차원

Nearest points in the original space:

herbal	0.320
moss	0.384
citrus	0.401
floral	0.412
fruity	0.462

Nearest points in the original space:

herbal	0.605
moss	0.687
citrus	0.761
whiteflower	0.844
floral	0.855



지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색 특징 향 최종



향 관련 유사도

Main1	Main2	Main3	Mid1	Mid2	Mid3
floral					
...	...	...	...	...	...

Word	1	2	...	19	20
floral	0.5560168	-0.5489983		0.18829721	-0.318604

향별 임베딩 벡터 값이 존재



지난주

데이터 수정

유사도

추천시스템

한계/의의

## 03. 유사도

색 특징 향 최종



최종 유사도



색  
유사도 행렬

특징  
유사도 행렬

향  
유사도 행렬

가중평균

최종  
유사도 행렬



지난주

데이터 수정

유사도

추천시스템

한계/의의

## 04. 추천시스템



추천시스템

우리가 선택한 방식은 무엇?

✓ *Contents-based Filtering*

상품 자체를 분석하여 비슷한 상품을 추천하는 방식!

사용자 데이터를 구할 수 없었습니다...

사용자가 특정 상품을 좋아하면, 이와 비슷한 다른 상품을 추천!



## 04. 추천시스템

지난주

데이터 수정

유사도

추천시스템

한계/의의

```
*Python 3.8.5 Shell*
Type "help", "copyright", "credits" or "license()" for more information.
>>>
===== RESTART: /Users/kimminjung/Desktop/final_perfume.py =====
1.여성 2.남성 3.유니섹스
어떤 성별의 향수를 원하시나요?
1
1.warm : 따뜻하고 여유롭고 풍요로운 향
2.green : 풀, 나뭇잎, 이끼를 연상시키는 향
3.herbal : 시원하고 상쾌하며 풀향이 나는 허브향
4.unusual : 파우더향이나 흙향, 금속의 향, 아스팔트의 향과 같이 향수의 구성에서 쓰일 수 있는 독특한 향
5.gourmand : 미각을 자극하는 맛있는 향
6.floral : 화사하고 부드러운 꽃 향
7.spicy : 다양한 향신료가 가지고 있는 자극적이면서도 따스하고 감성적인 향
8.aqua : 시원하고 투명한 물을 연상시키는 향
9.fruity : 상큼하고 달콤한 과일향
10.woody : 나무의 침착하고 깊은 향
어떤 느낌의 향을 좋아하시나요?
좋아하는 향 두 개를 골라주세요! ex)1,2
1,2
1.amber : 호박(보석)에서 느낄 수 있는 따뜻하고 스윗한 향
2.green : 풀, 나뭇잎, 이끼를 연상시키는 향
3.herbal : 시원하고 상쾌하며 풀향이 나는 허브향
4.unusual : 파우더향이나 흙향, 금속의 향, 아스팔트의 향과 같이 향수의 구성에서 쓰일 수 있는 독특한 향
5.gourmand : 미각을 자극하는 맛있는 향
6.floral : 화사하고 부드러운 꽃 향
7.spicy : 다양한 향신료가 가지고 있는 자극적이면서도 따스하고 감성적인 향
8.aqua : 시원하고 투명한 물을 연상시키는 향
9.fruity : 상큼하고 달콤한 과일향
10.woody : 나무의 침착하고 깊은 향
11.balsam : 나무의 수지 또는 바닐라와 같은 달콤하고 숙성되어 풍성한 감성의 향
12.musky : 모든 향을 부드럽게 감싸안은 관능적이고 따뜻한 향
13.beverage : 음료 레시피의 일부를 구성하는 재료를 활용하여 식욕을 돋우고 미뢰를 유혹하는 향
14.citrus : 레몬, 오렌지, 자몽 같은 가볍고 상큼한 감귤류 향
15.mossy : 나무 표면에 낀 이끼에서 느껴는 향으로 전체적으로 그늘린 듯한 텅텅한 향미가 은은한 향
16.white flower : 백합, 은방울꽃처럼 단아하고 부드러운 꽃 향
17.animalic : 풍부하고 따스한 분위기의 향
18.nutty : 아몬드, 땅콩, 헤이즐넛과 같은 견과류의 향으로 진하고 고소한 향
19.sweet : 달콤하고 먹음직스러운 향으로 후각적 자극을 미각으로 바꾸는 듯한 향
싫어하는 향기노트 1개를 골라주세요.

Ln: 23 Col: 29
```

향에 대한 설명을 제시하여

향.알.못들도

자신이 원하는 향수를

추천 받을 수 있는 시스템!



지난주

데이터 수정

유사도

추천시스템

한계/의의

## 05. 한계/의의



의의

1

정형데이터, 텍스트, 이미지(일부)등 다양한 데이터를 다룸

2

R과 파이썬을 모두 사용하여 진행함

#부싷다...

3

향수라는 객관화하기 힘든 제품을 다룸 #추상적인\_것도\_분석해봤다!