ScentMate
Fragrance Recommendation Systems

Kelly Peng
09/07/2017

Motivation

How to choose amongst $20{,}000$ perfumes?

ScentMate, find your perfect match.

What do I want to do

CONSUMER GROUPS

NEW USERS

Don't know what to wear

INTERMEDIATE USERS

Look for perfumes similar with what they liked before
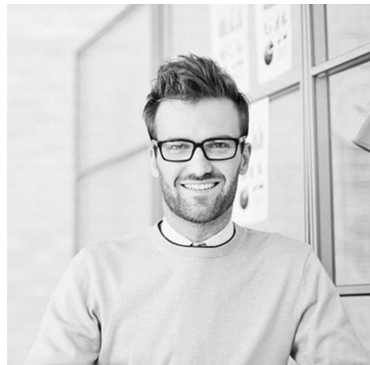
PERFUME LOVERS

Rated more than 3 perfumes on our website

What do I want to do

NEW USERS

INTERMEDIATE USERS

PERFUME LOVERS

CONSUMER GROUPS

Don't know what to wear

Look for perfumes similar with what they liked before

Rated more than 3 perfumes on our website

HOW TO SERVE THEM

Recommend based on quiz results

Recommend similar products

Recommend based on rating history

Model 1: Content–Based Model

Model 2: Collaborative Filtering

How to get there

Store scraped data in MongoDB on AWS EC2 instance

Develop a web app using Flask, Bootstrap, HTML, CSS

Parallel scrape data from website using 6 AWS EC2 instances

Data analysis, feature engineering, modeling, cross validation

Deploy everything onto AWS EC2 instance, run flask on AWS

The Data

Perfume Data
- Number of Perfumes: 20k
  - Brand: 1,824
  - Gender: 3
  - Note: 653
  - Theme: 30

User Rating Data
- Number of Ratings: 40k
  - Perfume ID
  - User ID
  - Rating Score: Range 2 – 10
  - User Comment

Model 1: Content-Based Recommender

Apply NMF and LDA to
user comments, choose 12
topics generated by LDA

Model 1: Content-Based Recommender

Apply NMF and LDA to user comments, choose 12 topics generated by LDA

12 Latent Topics found by LDA (Latent Dirichlet Allocation)

1. sweet, vanilla, apricot, chocolate, patchouli
2. gourmand, sweet, warm, heavy, winter
3. fig, fresh, green, plants, summer
4. gentle, elegant, mature, sexy, feminine
5. fresh, bergamot, citrus, neroli, amber
6. white flowers, fresh, quiet, soft, tea
7. young, fruity, sweet, lovely, lively
8. fresh, clean, summer, cool, cozy
9. rose, gentle, young, peony, feminine
10. spicy, warm, masculine, gentle, mature
11. oriental, mysterious, spicy, deep, dark
12. classic, chanel, dior, luxury, street

*Keywords are manually translated from Chinese text, thus no stemming or lemmatizing

Model 1: Content-Based Recommender

Apply NMF and LDA to
user comments, choose 12
topics generated by LDA

Combine LDA keywords
with other perfume features,
build perfume matrix

Model 1: Content-Based Recommender

Apply NMF and LDA to
user comments, choose 12
topics generated by LDA

Build content–based similarity
model based on Jaccard similarity

Combine LDA keywords
with other perfume features,
build perfume matrix

Model 2: Collaborative Filtering Recommender

	Base Model	**Item–item Similarity**	**Funk SVD**
Methodology	Predict mean for everything	Recommend based on the most similar items found by user ratings	Decompose utility matrix into two matrices with latent factors
Performance on Test Set	RMSE: 2.20	RMSE: 7.32	RMSE: 1.95
Reason		Utility matrix too sparse lead to little information going into each prediction	Should perform better if more data available

Model 2: Collaborative Filtering Recommender

Best!

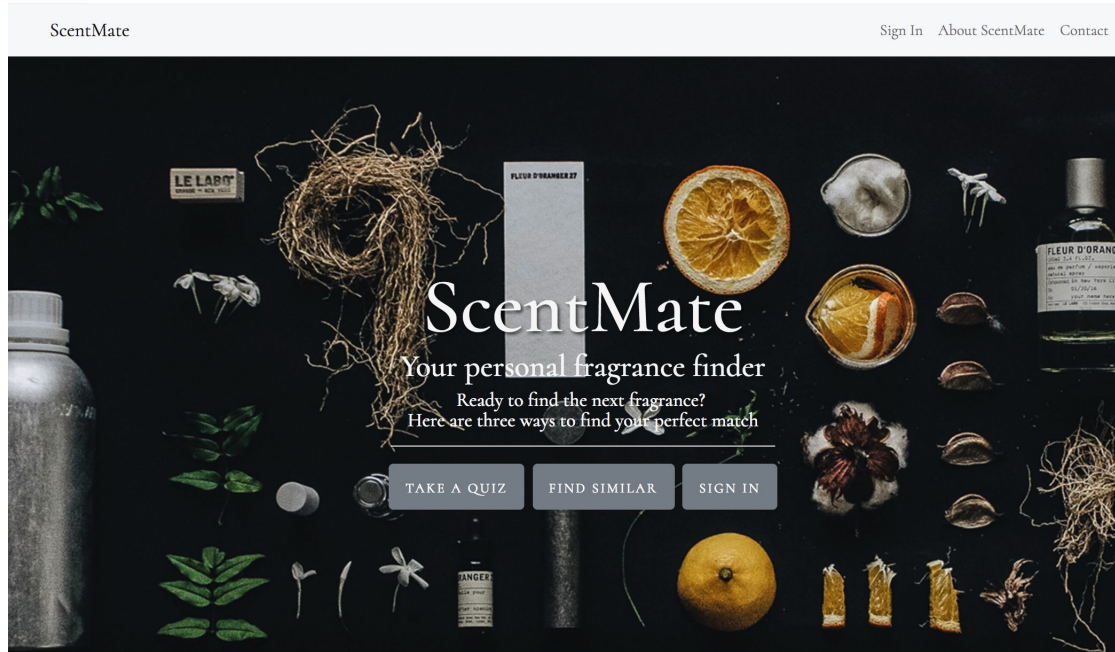|	**Base Model**	**Item–item Similarity**	**Funk SVD**
Methodology	Predict mean for everything	Recommend based on the most similar items found by user ratings	Decompose utility matrix into two matrices with latent factors
Performance on Test Set	RMSE: 2.20	RMSE: 7.32	RMSE: 1.95
Reason		Utility matrix too sparse lead to little information going into each prediction	Should perform better if more data available

Final model:
14 latent factors
Added regularization

Final Product: Web App
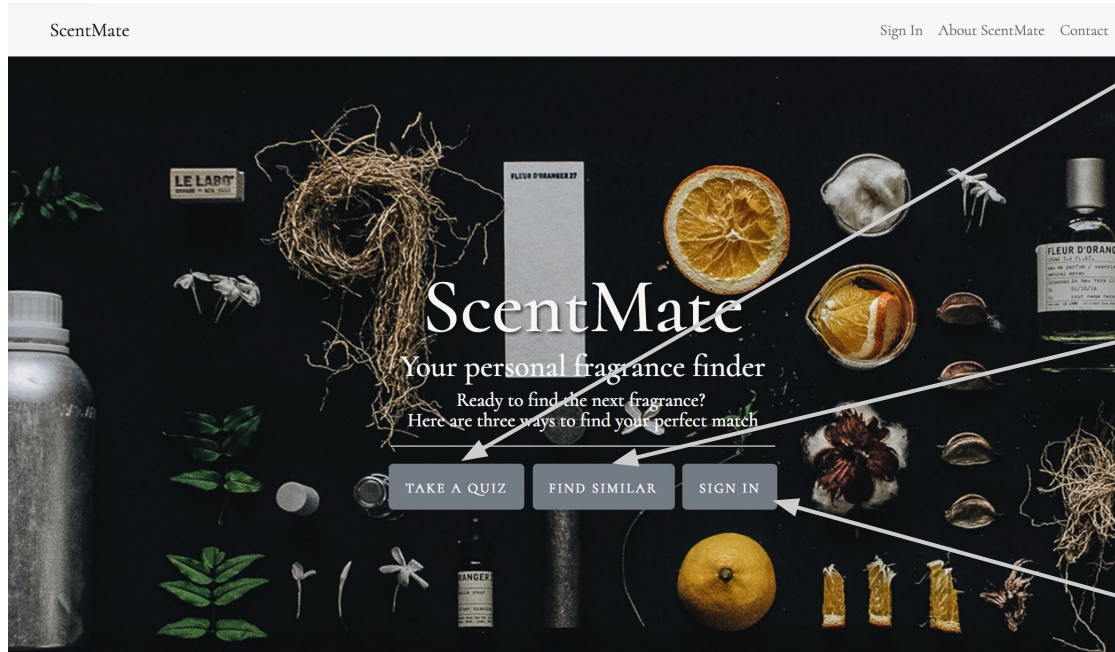
(Have a try on: 34.212.206.24)

ScentMate   Sign In   About ScentMate   Contact

ScentMate

Your personal fragrance finder

Ready to find the next fragrance?
Here are three ways to find your perfect match

TAKE A QUIZ   FIND SIMILAR   SIGN IN

NEW USERS

INTERMEDIATE USERS

PERFUME LOVERS

Final Product: Web App

(Have a try on: 34.212.206.24)

ScentMate

Sign In   About ScentMate   Contact

ScentMate
Your personal fragrance finder
Ready to find the next fragrance?
Here are three ways to find your perfect match

TAKE A QUIZ   FIND SIMILAR   SIGN IN

NEW USERS

INTERMEDIATE USERS
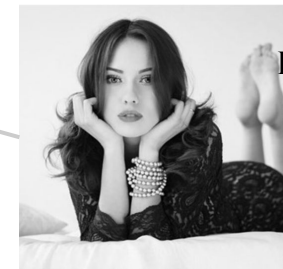
PERFUME LOVERS

Final Product: Web App
(Have a try on: 34.212.206.24)

ScentMate

Tell me something about you....

Your gender ○Female ○Male

How do you want to feel when wearing perfumes?
- ☐Sweet
- ☐Gentle
- ☐Elegant
- ☐Mature
- ☐Sexy
- ☐Feminine
- ☐Fresh
- ☐Classy
- ☐Soft
- ☐Easygoing
- ☐The girl next door
- ☐Young lady
- ☐Lovely
- ☐Lively

NEW USERS

Recommend based on quiz answers

Final Product: Web App

(Have a try on: 34.212.206.24)

ScentMate

Tell me somethin

Based on what you like, we found 5 perfumes match you the best:

Brand	Perfume Name	Gender	Theme	Notes
Hermes	爱马仕 橘绿之泉 Hermes Eau D`Orange Verte, 1979	Unisex	Floral Green	orange b
Kaloo	卡露儿 蓝熊宝宝 Kaloo Blue, 2001	Unisex	Floral	bergamo
Roger & Gallet	香邂格蕾 上善之水 Roger & Gallet Jean-Marie Farina Eau des Bienfaits	Female	Floral Fruity	blood or
Courreges	活希源 天空之水 Courreges La Fille de l'Air Eau d'Ete, 2016	Female	Floral	bergamo
Guerlain	娇兰 花草水语 - 初夏沁梨（夏日梨沙冰）Guerlain Aqua Allegoria Pera Granita, 2016	Female	Floral Fruity	grapefru musk, c
Kenzo	高田贤三 玫瑰黄 Kenzo Couleur Kenzo Jaune-Yellow, 2012	Female	Floral	bergamo

Your gender ○Female ○Male

How do you want to feel when wearing
- ☐Sweet
- ☐Gentle
- ☐Elegant
- ☐Mature
- ☐Sexy
- ☐Feminine
- ☐Fresh
- ☐Classy
- ☐Soft
- ☐Easygoing
- ☐The girl next door
- ☐Young lady
- ☐Lovely
- ☐Lively

NEW USERS

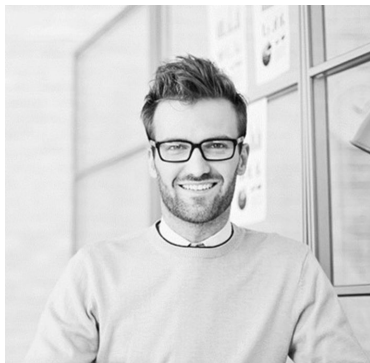Recommend based on quiz answers

*Result table has Chinese text because data is scraped from Chinese fragrance website

Final Product: Web App
(Have a try on: 34.212.206.24)

Intermediate Users

Recommend based on content similarity

I see, you like 迪奥 真我 Dior J`adore, 1999.

This is a Floral Fruity Female fragrance of Dior.

The key features of this fragrance are: magnolia, melon, peaches, pear, bergamot, oranges, late incense jade, plums, violet, orchids, freesia, jasmine, lily of the valley, rose, musk, v

When you are wearing this perfume, people will think of you as ...(need futther NLP work)

And because you like 迪奥 真我 Dior J`adore, 1999, we found some perfumes you will also like:

Brand	Perfume Name	Gender	Theme	Note
Lanvin	浪凡 我愿意（嫁给我） Lanvin Marry Me, 2010	Female	Floral Fruity	bitter orange, freesia, peaches, jasmine, rose, magnolia, musk, cedar, amber
Laura Biagiotti	罗娜 同名女士香水 Laura Biagiotti Laura, 1994	Female	Floral	lychee, plums, green leaves, peaches, bergamot, watermelon, Brazilian mahogany, c the valley, rose, sandalwood, amber, musk, vanilla, vetiver grass, cedar
YSL Yves Saint Laurent	圣罗兰 唯爱（反之亦然） YSL Yves Saint Laurent Vice Versa, 1999	Female	Floral Fruity	oranges, raspberry, blueberries, tomato, jasmine, orange blossom, late incense jade, iris, cedar, amber, vanilla
Ted Lapidus	泰德·拉皮迪斯 幻想 Ted Lapidus Fantasme, 1992	Female	Floral Fruity	pineapple, green leaves, violet, peaches, bergamot, rose, raspberry, iris root, jasmine
Clinique	倩碧 快乐 Clinique Happy, 1998	Female	Floral Fruity	oranges, oranges, plums, grapefruit, apple, bergamot, orchids, freesia, lily of the val

Final Product: Web App
(Have a try on: 34.212.206.24)

Based on what you like, we found 5 perfumes match you the best:

Brand	Perfume Name	Gender	Theme	Notes
Hermes	爱马仕 橘绿之泉 Hermes Eau D`Orange Verte, 1979	Unisex	Floral Green	orange blossom, oranges, jasmine, bergamot, lemon
Kaloo	卡露儿 蓝熊宝宝 Kaloo Blue, 2001	Unisex	Floral	bergamot, oranges, orange blossom, honeysuckle, musk
Roger & Gallet	香邂格蕾 上善之水 Roger & Gallet Jean-Marie Farina Eau des Bienfaits	Female	Floral Fruity	blood orange, oranges, bergamot, lemon, violet leaves, o
Courreges	活希源 天空之水 Courreges La Fille de l'Air Eau d'Ete, 2016	Female	Floral	bergamot, ozone, orange blossom, musk, cedar
Guerlain	娇兰 花草水语 - 初夏沁梨（夏日梨沙冰） Guerlain Aqua Allegoria Pera Granita, 2016	Female	Floral Fruity	grapefruit, bergamot, lemon, orange blossom, osmanthu musk, cedar
Kenzo	高田贤三 玫瑰黄 Kenzo Couleur Kenzo Jaune-Yellow, 2012	Female	Floral	bergamot, lemon, jasmine, orange blossom, vanilla, musl

Perfume Lovers

Recommend based on rating history

Future Work

1. Improve content-based recommender

   a. Price information not available, big drawback on model performance

   b. Word2Vec: combine multiple perfumes together to find a similar one

2. Build an automatic data processing pipeline

   a. Make sure data is not outdated

Thank you

Kelly Peng

kellypeng17@gmail.com

github.com/kellypeng

linkedin.com/in/pengyuwei

Appendix 1

Issues with data:

60% users rated < 3 perfumes
- Problem: Cannot cross-validate prediction
- Solution: Removed users with < 3 ratings

Number of ratings per user