A Content-based Skincare Product Recommendation System

Krystal Lee

Department of Computer Science, Earlham College, Richmond, IN

Motivation and Background

Hailed as the "fastest-growing category globally," skincare has taken over makeup with its increasing sale each year. In the past, consumers depended on best-sellers or in-store recommendations from the counter. However, everyone has different skin conditions, so these are not effective methods to judge compatibility between a product and a user.

To address this problem, this study designed a skincare product recommendation system based on the user's skin type and ingredient composition of a product.

- Content-based filtering takes into account the descriptions of the items as well as user preferences
- This method is suitable for making personalized recommendations and doesn't involve other users
- Short names CBF and IF-IPF were used to represent content-based filtering and ingredient frequency-inverse product frequency

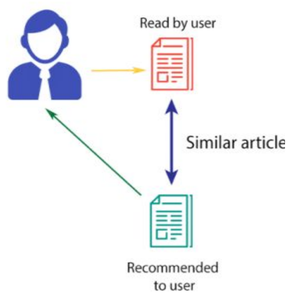

Figure 1: CBF diagram

Project Framework

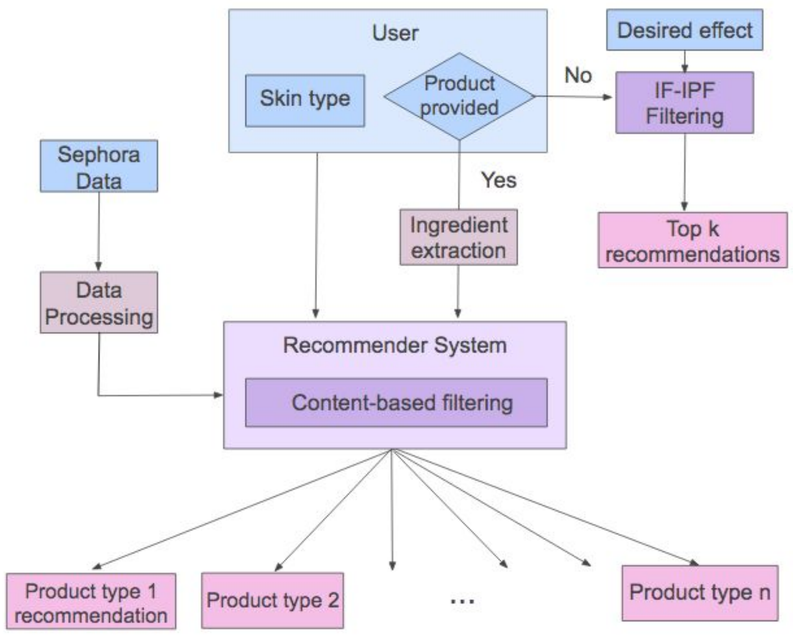

Figure 2: Project framework

Methods

Data Collection and Ingredient Extraction
An existing dataset on cosmetics from Jeong's project was used. Ingredients were tokenized to create a document term matrix that indicates the existence of ingredients in each product

	water	niacin	lanolin alcohol	Decyl oleate	serine	. . .	sh-polypeptide-1
Cosmetic 1	1	1	0	1	0	. . .	1

Figure 3: Cosmetic-ingredient matrix

Content-based Filtering
A user provides his or her skin type and selects a product from one of six categories (moisturizing cream, facial treatments, cleanser, face mask, eye treatment, and sun protection). The cosine similarity values are calculated for each product and ranked across all six categories.

```
Moisturizer
                                                Name      dist
0  Future Solution LX Total Regenerating Cream  0.991317
1       Benefiance WrinkleResist24 Night Cream  0.991129
2                  Vitamin C Glow Moisturizer  0.952674
3             Black Tea Firming Overnight Mask  0.873000
4                 GenOptics Spot Essence Serum  0.861700
```

Figure 4: Sample of content-based filtering result

IF-IPF Filtering
A user chooses his or her desired beauty effect from anti aging, moisturizing, oil control, acne treatment, redness control, and reduced pores along with skin type. Then, the Ingredient Frequency-Inverse Product Frequency values are obtained to identify the most important ingredient for each effect and the products containing it.

$$IF\text{-}IPF(i,X) = \sum_{p=1}^{m} \frac{n_p - \alpha_{p,i}}{n_p} \times log \frac{N}{pf(i)}$$

Equation 1: IF-IPF equation

```
Recommendations for Moisturizer

Crème de la Mer
The Moisturizing Soft Cream
Crème de la Mer Mini
```

Figure 5: Sample of IF-IPF filtering result

Validation

As part of the evaluation process, ratings of recommended products were filtered by the skin types of participants. Then, the number of reviews with recommended tags were divided by the total number of reviews as recorded in the table:

	Normal		Oily	
Product type	CBF (%)	IF-IPF (%)	CBF (%)	IF-IPF (%)
Moisturizer	86.96	72.00	85.71	80.47
Cleanser	77.14	85.11	52.94	85.65
Treatment	81.82	100.00	91.10	70.08
Face Mask	91.30	83.33	89.33	100.00
Eye Cream	75.53	81.25	76.19	93.13
Sun Protect	82.76	73.91	70.00	52.78

Table 1: Validation result for normal and oily skin

For participants with oily skin, CBF had 77.55% accuracy on average while IF-IPF generated 80.35% accuracy. For normal skin, CBF gave 82.59% while IF-IPF had 82.60%.

Outliers were formed due to low cost-effectiveness of some products. For both normal and oily skin types, the differences in percentages were less than 0.1.

Thus, it can be concluded that both methods are about equally efficient, and one can choose to use either method as needed.

Future Work

In the future, one could improve the system by incorporating brand preferences or price while making recommendations. With an appropriate data set, one could also try to implement the hybrid recommender system.

Acknowledgements

I would like to thank Dr. Xunfei Jiang for guiding me in all aspects of the project. I would also like to thank students who provided inputs for the filtering methods.

References

This poster is based on work described in "A Content-based Skincare Product Recommendation System", Gyeongeun Lee, available at https://portfolios.cs.earlham.edu/index.php/author/klee16/.