A Content-based Skincare Product Recommendation System

Gyeongeun Lee
Department of Computer Science
Earlham College
Richmond, Indiana, 47374
klee16@earlham.edu

ABSTRACT

In recent years, consumer interest in cosmetics has been increasing globally with a focus on skincare. In the past, consumers have depended on best-seller products 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. This proposal focuses on designing a skincare product recommendation system based on the user's skin type and ingredient composition of a product. Content-based filtering is used to identify the chemical components of products and find products with similar ingredient compositions. This method also allows users to input their desired beauty effect instead of a product name if they lack knowledge or have not found a product they like.

KEYWORDS

Content-based filtering, Recommender system, Cosmetics

1 INTRODUCTION

Hailed as the "fastest-growing category globally," skincare has taken over makeup with its increasing sale each year [15]. According to Trefis, a financial research and analysis firm, sales of skincare products in the U.S. surged by 13% in 2018, while makeup sales only grew by 1% [2]. Trefis also estimates the global skincare market to reach $180 billion, an increase of over 30% from the current stage, in the next five years. This growth is mainly due to the customers' pursuit of natural beauty as well as men's increased interest in skincare products. Companies are also adding anti-aging products for women to keep them as their primary consumers as they get older.

The need for advanced technology accompanied such growth as more customers started visiting the cosmetics counter to get product recommendations [11]. However, this process is often ineffective and time-consuming. The overwhelming quantity of accessible online information has also made it difficult for users to make correct choices [1]. The abundance of product information and reviews are perceived to be valuable. But at the same time, it prevents users from picking out desired information and making decisions based on their needs. Such difficulty has evoked the pressing need for personalized systems that could ease the access of data.

Researchers have proposed different recommender systems in an attempt to resolve the information overloading problem and facilitate the selection process [5]. The two most commonly adopted methods are collaborative filtering and content-based filtering. Recently, a hybrid approach that combines the two techniques was introduced in an attempt to maximize the benefits of both methods while covering their weaknesses.

However, it is still unclear which technique best measures the suitability of products for each customer. In fact, lots of online cosmetics stores still recommend bestsellers to customers regardless of their individual skin conditions [13]. Hence, there is a need for further investigation and improvement in the recommender systems for personal care products.

This proposal presents a content-based recommendation method that evaluates the similarity of ingredient composition within products. Instead of recommending within the same category, the new system recommends products across different categories to allow more effective recommendations. It also gives an option for the user to provide minimum input to get suggestions for skincare products. The system will be validated by taking the ratings of each product and comparing them to the results.

The paper starts by discussing related work and introducing existing recommender systems including collaborative filtering, content-based filtering, and hybrid filtering. Then, the design of the proposed system is explained in detail with a diagram. It not only elaborates on the framework but also covers each component of the recommender system. Moreover, the verification method for the system and test plans are added to the design section, followed by budget and timeline sections.

2 RELATED WORK

This section discusses existing recommendation methods, including collaborative filtering, content-based filtering, and a hybrid approach that mixes both. It focuses on the details of each technique, motivation, contribution, and the experimental or theoretical nature of the research. Moreover, the analysis and comparison of different algorithms used within each recommender system are presented with their results.

2.1 Collaborative Filtering

Collaborative filters use the information provided by users, such as clicks, likes, purchases, etc. Although they face a cold-start problem, they work well with enough amount of behavioral data [3].

Researchers who use collaborative filtering believe that identifying similar users can help match suitable products with the customer.

Matsunami et al. and Okuda et al. both adopted user similarity calculating method and analyzed reviews of cosmetic items [9, 10]. They used automatic scoring and k-means clustering to extract not only ratings but also textual reviews that contain individual preference and opinion. Ye also used collaborative filtering but focused on improving the weakness of the traditional method [16]. She attempted to alleviate the data sparsity problem and personalize the technique by making it more item-based. Matsunami et al. and Okuda et al. depended heavily on users' reviews to perform textual analyses while Ye focused more on the aspects of the items instead of users to recommend products.

The new content-based recommender proposed in the later section is similar to the research of Matsunami et al. and Okuda et al. in that it incorporates ratings of users. Although it does not filter products based on ratings, it uses them to validate the results. Like in Ye's research, the system not only considers users' opinions but also items' properties by taking their ingredients and comparing them with others.

2.2 Content-based Filtering

Another standard recommendation method is content-based filtering, which takes into account the descriptions of the items as well as user preferences [12]. Content-based filters tend to have an overspecialization problem, which is if someone is buying a mouse, the system will likely miss out on recommending a mouse pad or a keyboard.

Putriany et al. felt the need to personalize the skincare product recommendation as much as possible and adopted content-based filtering for their research [12]. The system was based on the items that were rated, liked, or chosen in the past by a particular user. Patty et al., like Putriany et al., focused on targeting the user profile but included factors such as cosmetic type, skin type, usage, price, description, and pictures for enhanced recommendation [11]. This approach can help personalize the method by going beyond their purchase habits. Unlike Putriany et al., Sato et al. tried to grasp other users' influence while using content-based filtering [14]. However, content-based filtering works only for active users, and it is hard to give a good recommendation if the information is not easy to categorize [12].

Jeong addressed this problem in her project and designed a recommender system based on the ingredients of the cosmetics [6]. She assumed that recommending skincare products should be separated from recommending movies because skin types and features of a person are more complex and sensitive. Honma et al., like Jeong, also took an approach to relate skin types to the ingredients of cosmetics [4]. Their method can overcome the weakness of the ones in collaborative filtering whose performance degrade as the number of ratings increases [12].

For algorithms, Putriany et al. used K-Means clustering to minimize the cluster performance index, square error, and error criterion [12]. They clustered different skincare products into five clusters based on the skin types and incorporated the users' likes on the products. As a result, they created an entity-relationship diagram along with a web-based system. They evaluated their method using purity, which measures the extent to which clusters contain a single class. The purity obtained was 0.29, which is not so high considering that the value 1 indicates 100% accuracy.

Patty et al. introduced content-based filtering using TF-IDF, which finds term relationships and ensures that the products conform to required specifications [11]. They used both the user profiles and item profiles to rank 40 products they selected for the test. They first divided the products into 17 different categories using the degree of similarity between the cosmetics data and user input. The item profiles included the cosmetic type and usage, and user profiles had skin type and price. Then they used cosine similarity to rank the items in descending order, which matched the results on their planning application recommendation. They did not provide any numerical measure for accuracy.

Jeong mainly used Natural Language Processing (NLP) for ingredients to match it with skin types [6]. She scraped 1472 items from Sephora with their brand, price, rating, ingredients, and appropriate skin types and preprocessed them. She divided the types of products into six and skin types into five and assigned binary values for ingredients present in each product. She then used cosine similarity to suggest the top 5 items with similar properties.

Similarly, Honma et al. used term frequencyinverse document frequency (TF-IDF) for their content-based filtering method [4]. They first defined user categories with similar skin quality, extracted evaluation of skin lotion, acquired ingredients for each skin lotion group and used TF-IDF to build equations to recommend skin lotion containing ingredients with high beauty effects. To do this, they considered 7 degrees of satisfaction, 15 types of effects (anti-aging, moisture, acne, etc.), reviews, and ingredients of skin products. At the end of their research, the number of invalidated products for each recommended product group was less than 5 percent, showing that their method is reliable.

In this group, Sato et al. focused on user preference and recommended items based on their likes, while other researchers incorporated skin type and personal user attributes. Despite varying algorithms, the trend of the use of content-based filtering was prioritizing user profiles over past purchase history. Honma et al. and Jeong's research both concerned ingredient composition within products but Honma et al. used reviews and checked for users' desired beauty effects while Jeong did not incorporate ratings. Honma et al.'s method could further personalize the recommendation, but will not work well with sparse and inconsistent reviews.

To overcome the overspecialization problem of content-based filtering, the newly proposed method closely follows Jeong's work and recommends items that are not limited to a single category. Instead, it gives recommendations across all 6 categories so that users do not have to search within each cosmetic type. One drawback of Putriany et al. and Jeong's research is that their system expects users to select an initial product as an input. But what if the user struggles with the selection process? The proposed method gives an option for the user not to have to input their product. Instead, it takes other desired information from the user to provide non content-based recommendations. Another drawback of Jeong's research is that it does not have any measure of verification for the system. Hence, the newly proposed method includes experiments at the end to test its reliability.

2.3 Hybrid Approach

Noticing the problems in the traditional methods, companies like Netflix and Google started adopting hybrid recommender systems [7]. Experiments on the live traffic of the website done by Google suggest that the hybrid method improves the quality of recommendation [8]. With similar assumptions, some researchers have used hybrid filtering to maximize the benefits of both collaborative and content-based filtering.

Hansson proposed a hybrid recommender for online products using k-means++ [3]. Using the data set from an online book retailer and fashion retailer, she obtained the values for precision and recall for each method tested on both data sets. She concluded that her algorithms do not have the same functionality across different data sets, and combinations of strong algorithms do not produce better results. James and Rajkumar also proposed a hybrid method and added the time sequence method for collaborative filtering [5]. Unlike Hansson's, their research is theory-based and was not tested on an actual data set. They suggested three different directions for their algorithm: item similarity, bipartite projection, and spanning tree. Since the time sequence model learns the change in data over time, the authors concluded that it will generate higher accuracy by using static data.

Although a hybrid approach seems to have potential in the skincare domain, it requires a data set that involves both the behavioral information of the user as well as the product information. Such data set, however, is scarce in skincare, so the proposed method only includes content-based filtering.

3 DESIGN AND IMPLEMENTATION

As illustrated in Figure 1, the proposed system offers content-based filtering or non content-based filtering depending on if the user inputs a product in the beginning.
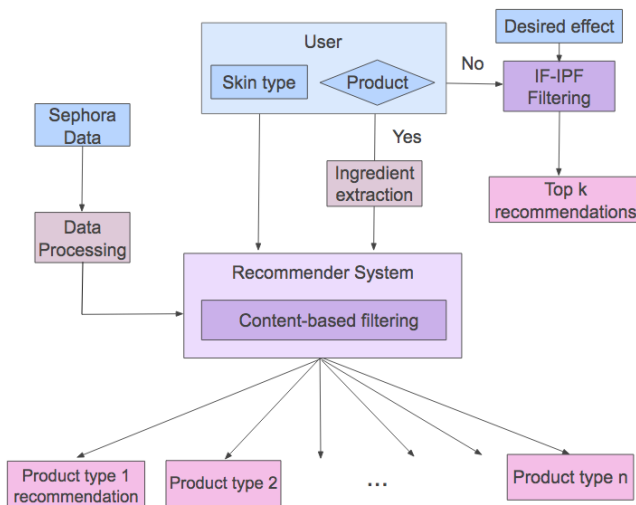

Figure 1: Framework of the content-based recommender system

Content-based Filtering. A user provides one of five skin types (combination, dry, normal, oily, and sensitive) and selects a product from one of six categories (moisturizing cream, facial treatments, cleanser, face mask, eye treatment, and sun protection). The skin type directly maps to the recommender system while ingredients are extracted from the product. Then, the skin type of the user and ingredients of the product are sent to the content-based recommender system along with Sephora data, which contains information about other products. In this method, recommendations of products are made across all six categories. After evaluating the similarity of ingredient composition within products, the system returns k number of recommendations for each of the n product types.

IF-IPF Filtering. If a user provides his or her skin type but not the product, the system takes their input for one desired beauty effect among anti-aging, moisturizing, oil control, acne treatment, redness control, and reduced pores. Then, it uses term frequency-inverse document frequency (TF-IDF) to filter products and makes recommendations in a similar format to content-based filtering.

3.1 Data Collection

An existing data set on cosmetics from Jeong's research was used in this project [6]. The data was scraped from sephora.com, a website that offers beauty products from multiple brands. Among many categories of personal care items, only six were extracted to focus on skincare products. These six categories include moisturizing cream, facial treatments, cleanser, facial mask, eye treatment, and sun protection. The data set consists of 1472 items which includes information about the brand, name, price, rank, skin types, and chemical components of each product.

Additionally, star ratings for all 1472 items will be extracted from sephora.com along with the reviewers' skin types. The extraction will be done using a tool called Scrapestorm[1] that allows data mining from different websites. This data set will be used specifically to evaluate the efficiency of this method after the implementation of the content-based recommender system. The design of the validation method will be further discussed in Section 3.5.

3.2 Ingredient Extraction

The ingredient extraction method closely follows Jeong's approach [6]. Initially, the collected data are filtered by the skin type input of the user. Once the user selects a product of their choice, the system extracts its ingredients and sends it to the recommender system along with the Sephora data set. The list of all ingredients are taken from the ingredients column of the data set and split into tokens. Once the duplication is checked, each chemical element is be given a unique index to be stored in a dictionary.

Next, the document term matrix (DTM) is created between the products and corresponding ingredients for each product. An empty matrix is initialized and filled with zeros. Here, the number of rows represents the number of skincare products, and the number of columns represents the number of total ingredients. Then, one-hot encoding is used to fill in the cosmetic-ingredient matrix with either 1 (present) or 0 (not present), depending on the existence of

---

[1]scrapestorm.com

ingredients in each product. An example of such matrix is illustrated in Figure 2.

Figure 2: Cosmetic-ingredient matrix [6]

3.3 Content-based Filtering

Once the ingredients are extracted and processed, they are passed into the recommender system along with the user's skin type. Content-based filtering is used based on Jeong's project [6]. In this method, cosine similarity is adopted to measure the similarity of the ingredient composition between products. It is applied to produce k recommendations for n product categories to rank cosmetics that have similar properties with the original product.

$$Cos\theta = \frac{\vec{a} \cdot \vec{b}}{||\vec{a}|| \, ||\vec{b}||} = \frac{\sum_1^n a_i b_i}{\sqrt{\sum_1^n a_i^2} \sqrt{\sum_1^n b_i^2}} \tag{1}$$

Using the matrix generated in Section 3.2, all cosmetic items are vectorized into two-dimensional coordinates. These coordinates are plugged into equation 1 to obtain the distances between different points. Finally, these values are arranged in ascending order to rank from the most similar to the least similar products. The process is repeated by passing in different product categories to filter the data. Splitting the data set into different types allows the system to recommend products across multiple categories.

3.4 IF-IPF Filtering

In case the users have never used or liked any product from Sephora, they can receive recommendations based on their skin type and desired beauty effects. The top ingredient that boost certain beauty effect is identified by calculating TF-IDF values which, in my project, is called ingredient frequency-inverse product frequency (IF-IPF). The IF-IPF values are derived using equations 2, 3, and 4 [4].

$$IF(i, X) = \sum_{p=1}^{m} \frac{n_p - \alpha_{p,i}}{n_p} \tag{2}$$

n_p: the number of unique ingredients included in product p in beauty effect group X
m: the number of products in beauty effect group X
$\alpha_{p,i}$: the rank of ingredient i listed in product p

$$IPF(i) = \log \frac{N}{pf(i)} \tag{3}$$

N: the number of products in the data set
$pf(i)$: the number of products including ingredient i

$$IF\text{-}IPF(i, X) = IF(i, X) \times IPF(i) \tag{4}$$

The products that contain those ingredients are then filtered and sorted across 6 product categories. Finally, the top k recommendations from each category are returned to the user.

3.5 Test Plan

To evaluate the performance of the content-based recommender system, ratings of users filtered by the inputted skin type were extracted from Sephora using Scrapestorm. Some ratings had "recommends this product" tags while others did not. These tags are important indicators of whether or not users were satisfied with the product. The reviews with the tags were labeled 'yes' and others were left blank. The number of tagged reviews were then counted and divided by the total number of reviews. The percentages of the result were derived to judge the effectiveness of the recommendation system.

3.6 Results

Content-based Filtering. A group of 5 female students, each with distinct skin type, were asked to identify their skin type. To verify the result, an online quiz from Interact[2] was used. Next, they were asked to provide a product they have used and liked on sephora.com. These inputs were put into the content-based recommender system, and five recommendations were generated for all 6 product categories as shown in Figure 3.

Figure 3: Sample of content-based filtering result

Given the large number of reviews per product, validation was done on two female students who have normal and oily skin. The result is shown in Table 1.

Product type	Normal		Oily	
	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

IF-IPF Filtering. 5 Male students with minimal knowledge of cosmetic products were asked to provide their skin type along with their desired effect. They covered four out of five available types. Again, these were either found or verified using the online quiz. They received recommendations like in Figure 3 based on their inputs. For validation, recommendations of two male students with normal and oily skin were used. The result is also shown in Table 1.

In Table 1, short names CBF and IF-IPF were used to represent content-based filtering and ingredient frequency-inverse document frequency. For participants with normal skin, CBF gave 82.59% accuracy rounded to the second decimal place while IF-IPF had 82.60%. In the case of oily skin, CBF had 77.55% on average while IF-IPF generated 80.35% accuracy. Outliers such as 52.94% were formed due to low cost-effectiveness of expensive skincare products. But other rates were relatively consistent and above 70%, indicating that the recommendations were reasonable. Furthermore, the differences in percentages between CBF and IF-IPF were less than 0.1 for both normal and oily skin types. Thus, it can be concluded that both methods are about equally efficient, and one can choose to use either method as needed. It is also important to note that the number of reviews were not consistent, ranging from 20 to more than 100. Since they were not evenly weighted, the accuracy of the result might have lowered.

CONCLUSION

This project implemented content-based filtering and IF-IPF filtering to make personalized recommendations based on user profiles and preferences. IF-IPF performed slightly better in the two test cases, but only by less than 0.1%. Both filtering methods produced accuracy higher than 75% on average, so they are both effective ways of recommending suitable products to users, given that 50% is neutral. But again, the result comes from just a small sample, and therefore further investigation is needed to obtain more reliable accuracy. 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.

ACKNOWLEDGEMENT

REFERENCES

[1] Shlomo Berkovsky and Jill Freyne. 2015. Web Personalization and Recommender Systems. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Aug. 2015). https://doi.org/10.1145/2783258.2789995

[2] Ahiza Garcia. 2019. The skincare industry is booming, fueled by informed consumers and social media. https://www.cnn.com/2019/05/10/business/skincare-industry-trends-beauty-social-media/index.html

[3] Linda Hansson. 2015. *Product Recommendations in E-commerce Systems using Content-based Clustering and Collaborative Filtering.* Master's thesis. Lund University, Lund, Sweden.

[4] Hirotoshi Honma, Yoko Nakajima, and Haruka Aoshima. [n.d.]. *Recommender System for Cosmetics Based on User Evaluation and Ingredients Information.* Technical Report. https://www.jtnews.jp

[5] Ndengabaganizi Tonny James and K. Rajkumar. 2017. Product Recommendation Systems based on Hybrid Approach Technology. *International Research Journal of Engineering and Technology* 4, 8 (Aug. 2017). www.irjet.net

[6] Jiwon Jeong. 2018. For Your Skin Beauty: Mapping Cosmetic Items with Bokeh. https://towardsdatascience.com/for-your-skin-beauty-mapping-cosmetic-items-with-bokeh-af7523ca68e5

[7] Kismet K. [n.d.]. Netflix: Recommendations Worth a Million. https://rpubs.com/kismetk/Netflix-recommendation

[8] Jiahui Liu, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized news recommendation based on click behavior. *Proceedings of the 15th international conference on Intelligent user interfaces* (Feb. 2010). https://doi.org/10.1145/1719970.1719976

[9] Yuki Matsunami, Asami Okuda, Mayumi Ueda, and Shinsuke Nakajima. 2017. User Similarity Calculating Method for Cosmetic Review Recommender System. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Vol. 1.

[10] Asami Okuda, Yuki Matsunami, Mayumi Ueda, and Shinsuke Nakajima. 2017. Finding Similar Users Based on Their Preferences against Cosmetic Item Clusters. (2017). https://doi.org/10.1145/3151759.3151829

[11] Joanna Cristy Patty, Elika Thea Kirana, and Made Sandra Diamond Khrismayanti Giri. 2018. Recommendations System for Purchase of Cosmetics Using Content-Based Filtering. *International Journal of Computer Engineering and Information Technology* 10, 1 (Jan. 2018), 1–5. www.ijceit.org

[12] Villia Putriany, Jaidan Jauhari, and Rahmat Izwan Heroza. 2019. Item Clustering as An Input for Skin Care Product Recommended System using Content Based Filtering. *Journal of Physics: Conference Series* (2019).

[13] X Ren. 2016. SKII Recommender System Design. http://xren615.github.io/post/sk2_rs/

[14] Tae Sato, Masanori Fujita, Minoru Kobayashi, and Koji Ito. 2013. Recommender System By Grasping Individual Preference and Influence from other users. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.*

[15] Sangeeta Singh-Kurtz. 2019. Luxury skincare is driving record profits in the beauty industry. https://qz.com/1544306/luxury-skincare-products-fuel-record-profits-in-beauty-industry/

[16] Hongwu Ye. 2011. A Personalized Collaborative Filtering Recommendation Using Association Rules Mining and Self-Organizing Map. *Journal of Software* 6, 4 (April 2011). https://doi.org/10.4304/jsw.6.4.732-739