OPEN ACCESS
Indo-JC
Indonesia Journal of Computing
ISSN 2460-9056
socj.telkomuniversity.ac.id/indojc

Ind. Journal on Computing
Vol. 9, Issue. 3, December 2024. pp. 220-231
doi:10.34818/indojc.2024.9.3.983

Evaluating Non-Negative Matrix Factorization and Singular Value Decomposition for Skincare Recommendation Systems

Ahmad Indra Nurfauzi [1], Agung Toto Wibowo [2]

School Of Computing, Telkom University
Telecommunication Street No. 1, Bandung 40257, Indonesia

[1] indranurfa@student.telkomuniversity.ac.id
[2] agungtoto@telkomuniversity.ac.id

Abstract

Facial skincare plays a crucial role in maintaining clean, healthy, and radiant skin. Recommendation systems, such as Collaborative Filtering and Content-Based Filtering, can help users discover suitable skincare products based on their preferences and reviews. This study compares two Matrix Factorization techniques Non-Negative Matrix Factorization (NMF) and Singular Value Decomposition (SVD) to enhance the accuracy and relevance of skincare product recommendations. The results reveal that the SVD model outperforms NMF, achieving a Mean Absolute Error (MAE) of 0.7190, Root Mean Squared Error (RMSE) of 1.0104, Precision of 0.8054, Recall of 0.8144, and an F-1 score of 0.8099. In contrast, the NMF model produced an MAE of 0.7074, RMSE of 1.1052, Precision of 0.7865, Recall of 0.7987, and an F-1 score of 0.7926. These findings demonstrate that both models provide accurate recommendations, with SVD offering more precise and relevant predictions for skincare product recommendations.

Keywords: Matrix Factorization, Collaborative Filtering, Skincare, Non-Negative Matrix Factorization, Singular Value Decomposition

Abstrak

Perawatan kulit wajah sangat penting untuk menjaga kulit tetap bersih, sehat, dan bercahaya. Sistem rekomendasi, seperti Collaborative Filtering dan Content-Based Filtering, dapat membantu pengguna menemukan produk perawatan kulit yang sesuai berdasarkan preferensi dan ulasan mereka. Penelitian ini membandingkan dua teknik Matrix Factorization, yaitu Non-Negative Matrix Factorization (NMF) dan Singular Value Decomposition (SVD), untuk meningkatkan akurasi dan relevansi rekomendasi produk perawatan kulit. Hasil penelitian menunjukkan bahwa model SVD mengungguli NMF, dengan Mean Absolute Error (MAE) sebesar 0,7190, Root Mean Squared Error (RMSE) 1,0104, Precision 0,8054, Recall 0,8144, dan F-1 score 0,8099. Sebaliknya, model NMF menghasilkan MAE 0,7074, RMSE 1,1052, Precision 0,7865, Recall 0,7987, dan F-1 score 0,7926. Temuan ini menunjukkan bahwa kedua model dapat memberikan rekomendasi yang akurat, dengan SVD memberikan prediksi yang lebih tepat dan relevan untuk rekomendasi produk perawatan kulit.

Kata Kunci: Matrix Factorization, Collaborative Filtering, Skincare, Non-Negative Matrix Factorization, Singular Value Decomposition

Received on August, 2024. Accepted on December, 2024

I. INTRODUCTION

THE facial skin is an essential part of the body that requires care to maintain a clean, healthy, and radiant appearance. Facial skin care, commonly referred to as skincare, involves activities designed to keep the skin in optimal condition—clean, healthy, and glowing [1]. Skincare is achieved through the use of various supportive products tailored to individual skin conditions. These products are incorporated into skincare routines that help maintain the skin's health, cleanliness, and care.

According to data from the International Market and Consumer Data Portal, Statista, in 2023 [2], the cosmetic industry market in Indonesia is expected to grow by 4.59 percent annually during the period of 2023-2028. This projection includes products such as skincare and personal care. Additionally, the National Food and Drug Agency (BPOM) reported an increase in the number of cosmetic industry players from 819 business units in 2021 to 913 business units in 2022, reflecting a growth of 20.6 percent.

The rising demand for skincare products in Indonesia is a clear reflection of a dynamic market, further fueled by the rapid growth of online retail platforms. Consumers today are presented with a vast array of options, ranging from moisturizers, facial cleansers, and toners to serums and other skincare products. While this variety offers more choices, it often leaves both beginners and experienced skincare enthusiasts grappling with decision fatigue. The challenge lies not only in selecting products that suit individual skin types but also in navigating the overwhelming number of brands available.

One possible solution to address this challenge is the implementation of a recommendation system. Such a system can help consumers navigate the extensive range of products by offering tailored suggestions based on user reviews, skin compatibility, and personalized preferences [3]. By connecting consumers with products that meet their unique needs, a recommendation system can enhance decision-making, leading to greater satisfaction and more effective skincare routines.

A Recommendation System is a system capable of suggesting items or content that are likely to be chosen, used, or purchased by users [4]. Recommendation systems are used by e-commerce sites to provide product suggestions to their customers. Products can be recommended based on overall top sales on the site, customer demographics, or analysis of past customer shopping behavior as a prediction for future shopping behavior.

In recommendation systems, there are two commonly used methods: Collaborative Filtering and Content-Based Filtering. In the collaborative filtering method, recommendations are made based on the estimated ratings of an item from other users who have similar preferences [5]. Content-Based Filtering is one of the earliest and most popular methods in recommendation systems. The principle of this method is to recommend objects that are similar to other objects that the user liked in the past. Similarity between objects is determined by the values of their characteristics [6].

The relationship between positive reviews of skincare products and ease of recommendation to users has become increasingly significant amidst the rapid influx of skincare products available today. As the variety of skincare products continues to grow, consumers often face challenges in selecting products that match their needs and preferences. In this context, Collaborative Filtering recommendation systems play a crucial role. By leveraging user reviews or product ratings, these systems effectively filter and recommend high-quality skincare products. Collaborative filtering's ability to analyze and combine historical rating data from various users ensures that consumers receive personalized and relevant recommendations. Thus, collaborative filtering provides an intelligent solution for navigating the diverse skincare product landscape, ensuring that positively reviewed products are easily accessible and enjoyed by consumers.

Non-negative Matrix Factorization (NMF) is one of the methods used in collaborative filtering. NMF is a matrix factorization method that decomposes a matrix into two smaller matrices. It operates under the assumption that all values in the factorized matrices are non-negative [7]. In recommendation systems, such matrices typically contain data such as user-provided ratings or reviews of products. In addition to NMF, Singular Value Decomposition (SVD) is another popular method in recommendation systems. SVD involves the decomposition of a matrix into three matrices. It gained prominence when Simon Funk introduced it during the Netflix Prize competition in 2007 [8].

II. LITERATURE REVIEW

Recommendation systems using matrix factorization have become an increasingly popular research topic in recent years. This research focuses on developing systems that can provide product or service recommendations tailored to individual preferences based on various factors such as user history, previous interactions, and rating patterns. In this context, several studies have explored the use of techniques like Singular Value Decomposition (SVD) and Non-Negative Matrix Factorization (NMF) to enhance recommendation accuracy.

The research conducted by Adyatma et al. (2023) [7] utilized a dataset from Goodreads, including book data and ratings data. In this study, Collaborative Filtering methods were used to compare two different algorithms: Singular Value Decomposition (SVD) and Alternating Least Squares (ALS). The results showed that SVD achieved better accuracy with an RMSE value of approximately 0.86822, an MAE value of approximately 0.6903, an F1-Score of approximately 0.827923, and a Precision of approximately 0.568347. In contrast, the ALS algorithm had an RMSE value of approximately 1.09320, an MAE value of approximately 0.86479, an F1-Score of approximately 0.000304, and a Precision of approximately 0.000596.

The research conducted by Nissa et al. (2023) [8] investigated a skincare recommendation system using Collaborative Filtering methods to compare two different algorithms: Singular Value Decomposition (SVD) and Alternating Least Squares (ALS). The results showed that the ALS method, compared to SVD, produced a higher RMSE value with 10-fold CV, which was 1.00949 compared to 1.00915.

The research conducted by Yoshua et al. (2021) [9] focused on a music recommendation system using Collaborative Filtering. This study evaluated the system using the metrics Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) to measure the performance of the recommendation system. The evaluation results indicated that the Singular Value Decomposition (SVD)++ algorithm provided the best performance with an RMSE of 0.0386 and an MAE of 0.0228. SVD also showed good performance with an RMSE of 0.0457 and an MAE of 0.0291. On the other hand, the Nearest Neighbors (KNN) algorithm using User-User and Item-Item methods with Cosine Similarity or Pearson Similarity showed uniform results with RMSE and MAE values of 0.1295 and 0.0526, respectively. The Probabilistic Matrix Factorization (PMF) and Non-Negative Matrix Factorization (NMF) algorithms showed RMSE and MAE values of 0.1090 and 0.0325, and 0.0677 and 0.0289, respectively.

A. Recommendation System

A recommendation system is a system designed to provide information and suggestions that help users make decisions based on existing data [10]. The goal of a recommendation system is to deliver effective and meaningful content (items) to active users on the platform [11]. These systems can employ various methods [12], which are utilized to recommend products, services, or information to users based on their preferences. The methods include Collaborative Filtering, Demographic Filtering, Content-Based Filtering, and Hybrid Filtering.

Content-Based Filtering generates recommendations based on the user's past choices [13]. This method also produces suggestions by analyzing the content of the items intended for recommendation. Demographic Filtering operates on the principle that individuals sharing similar personal attributes (e.g., gender, age, nationality) will likely have similar preferences [14].

Collaborative Filtering allows users to rate a set of items (e.g., videos, songs, movies) on a collaborative platform. Once sufficient information is stored in the system, recommendations can be generated for each user based on data provided by others with the most similar preferences [15]. Meanwhile, Hybrid Filtering combines two or more methods, such as collaborative filtering with demographic filtering or collaborative filtering with content-based filtering [16].

B. Collaborative Filtering

Collaborative Filtering is a popular recommendation algorithm that generates predictions and recommendations based on ratings or opinions from other users within the system [17]. The Collaborative Filtering method uses data based on similarities in consumer characteristics to provide information based on patterns from similar groups of consumers, thereby enabling the delivery of new information to consumers [18].

Collaborative Filtering systems, such as GroupLens [19], utilize user rating data to calculate similarities or weights between users or items. Based on these calculated similarity values, the system makes predictions or recommendations. Memory-based Collaborative Filtering, often referred to as CF, is particularly integrated into commercial platforms [20]. A well-known example of memory-based CF is the application of Collaborative Filtering to provide recommendations related to items.

Model-based CF techniques, on the other hand, use raw rating data to estimate or learn a model for making predictions [9]. This model can include data mining or machine learning algorithms. Notable model-based CF techniques involve Bayesian belief networks (BNs) CF models, clustering CF models, and latent semantic CF models [21].

C. Matrix Factorization

Matrix factorization operates by decomposing a large matrix into smaller matrices. The final result of this matrix is obtained from the dot product between the user matrix and the transpose of the item matrix. This method effectively captures latent features that describe both users and items, allowing the recommendation system to make accurate predictions even with sparse data [22]. By focusing on these latent factors, matrix factorization enables the system to generalize well across a wide range of products, enhancing the relevance of recommendations.

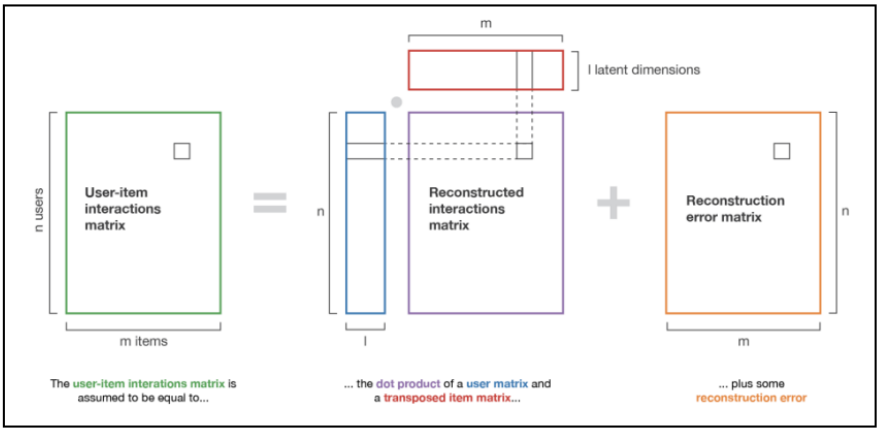

Fig 1.Illustration of the matrix factorization method

1. Non-negative Matrix Factorization (NMF)

Non-negative Matrix Factorization (NMF) is a matrix factorization method that decomposes a matrix into two smaller matrices. NMF is a matrix factorization technique that maintains the assumption that all values in the factor matrices are non-negative. This approach is particularly useful when the data is non-negative, such as in the case of user ratings for items. The prediction formula for NMF is given by equation (1):

$$\hat{r}_{ui} = q_i^T p_u \tag{1}$$

The optimization procedure is stochastic gradient descent (with regularization), with a specific step size selection that ensures the factors remain non-negative, provided the initial values are also positive. The formula for updating values using the regularized single-element-based NMF (RSNMF) model, based on the findings of Xin Luo et al. (2014) [23], is given by equation (2). At each step of the SGD procedure, the factors for users and items are updated according to this formula.

$$p_{uf} \leftarrow p_{uf} \cdot \frac{\sum_{i \in Iu} q_{if} \cdot r_{ui}}{\sum_{i \in Iu} qif \cdot \hat{r}_{ui} + \lambda_u |I_u| p_{uf}} \tag{2}$$

$$q_{if} \leftarrow q_{if} \cdot \frac{\sum_{i \in Ui} p_{uf} \cdot r_{ui}}{\sum_{u \in Ui} puf \cdot \hat{r}_{ui} + \lambda_i |U_i| q_{if}}$$

2. Singular Value Decomposition (SVD)

SVD is one of the popular matrix factorization methods used in recommendation systems. SVD is an algorithm developed by Simon Funk during the Netflix Prize competition in 2007 [24]. SVD is a matrix factorization technique that decomposes a matrix into three matrices. The matrix decomposition concept from the matrix factorization algorithm is applied by the SVD algorithm in the form of a formula to produce prediction values based on Ricci et al. (2011) [25]. The form of the prediction formula (3):

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T p_u \tag{3}$$

To estimate or compute all unknown rating values, SVD can achieve this by minimizing the regularized square error using the formula described in Formula (4):

$$\min_{p^*, q^*, b^*} \sum_{r_{ui} \in R_{train}} (r_{ui} - \hat{r}_{ui})^2 + \lambda(b_i^2 + b_u^2 + \|q_i\|^2 + \|p_u\|^2) \tag{4}$$

To reduce the errors produced by the algorithmic model in making predictions, steps are required that are performed by gradient descent using the formulas described in Formula (5):

$$\begin{gathered} b_u \leftarrow b_u + \gamma(\mathrm{e_{ui}} - \lambda \mathrm{b_u}) \\ b_i \leftarrow b_i + \gamma(\mathrm{e_{ui}} - \lambda \mathrm{b}_i) \\ p_u \leftarrow p_u + \gamma(\mathrm{e_{ui}} \cdot \mathrm{q_i} - \lambda \mathrm{p_u}) \\ q_i \leftarrow q_i + \gamma(\mathrm{e_{ui}} \cdot \mathrm{p_u} - \lambda \mathrm{q_i}) \end{gathered} \tag{5}$$

D. Skincare

In general, skincare is the activity of caring for the outer skin of the body using specific products to maintain skin health and appearance [26]. Skincare is not only for women but also used by men. Skincare products come in various types, but the most common include cleansers, facial washes, toners, moisturizers, serums, and sunscreens [27]. Facial wash functions as a facial cleanser and is available in various forms such as milk cleanser, balm cleanser, oil cleanser, and micellar water. Toner is used to refresh the skin and reduce excess oil on the face. Moisturizers are essential for maintaining skin hydration and preventing damage from makeup use and sun exposure. Sunscreen is also necessary to protect the skin from the harmful effects of UV rays from the sun [28].

E. Evaluation Metric

1. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is used to measure the difference as the average value between the algorithm's predictions and the actual ratings given by users [29]. MAE is calculated using equation (6):

$$MAE = \frac{\sum_i^k (p_i - r_i)}{k} \tag{6}$$

2. Root Mean Squeare Error (RMSE)

Root Mean Square Error (RMSE) is a metric used to measure the accuracy of a predictive model by calculating the square root of the average of the squared errors (the difference between the predicted and actual values) [29]. RMSE is calculated using equation (7):

$$RMSE = \sqrt{\frac{\sum_i^k (p_i - r_i)^2}{k}} \tag{7}$$

3. Precision

Precision is the ratio of the total number of relevant items to the total number of recommended items [29]. Precision can be calculated using equation (8):

$$Precision = \frac{|\ Interesting\ Items\ \cap Recommended\ Items\ |}{|\ Recomended\ Items\ |} \tag{8}$$

4. Recall

Recall is the ratio of the total number of relevant items that have been recommended to the total number of relevant items overall [29]. Recall is calculated using equation (9):

$$Recall = \frac{|\ Interesting\ Items\ \cap Recommended\ Items\ |}{|\ Interesting\ Items\ |} \tag{9}$$

5. F-1 Score

The F-1 Score is a combination of precision and recall, used to measure the balance between precision and recall, and is particularly useful when there is an uneven class distribution [29]. The F-1 Score is calculated using equation (10):

$$F1 = 2 \times \frac{precison \ \times recall}{precision + recall} \tag{10}$$

III. RESEARCH METHOD

In designing a skincare recommendation system using matrix factorization with NMF and SVD calculations, there are important components such as preprocessing, data splitting, and the computation using NMF and SVD algorithms. When designing the architecture of a system, a flowchart is needed to illustrate the steps of how the system operates. Fig. 2 shows the flowchart of the designed system:

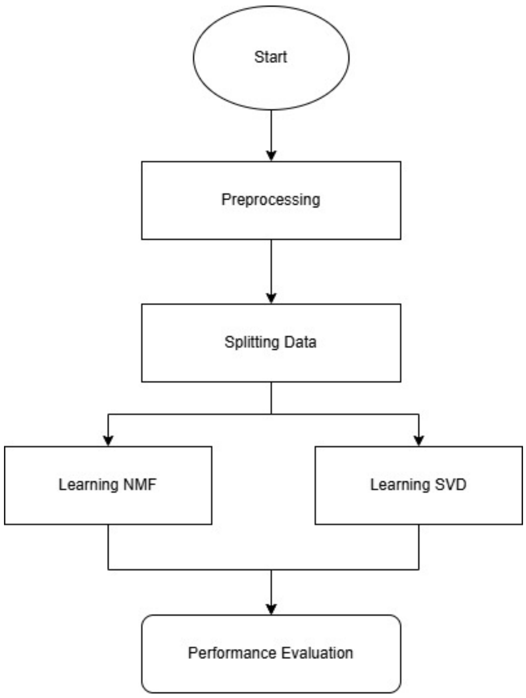

Fig 2. Flowchart Design System

A. Dataset

In this study, the dataset used is sourced from the Kaggle website, and the data was obtained from information available on the Sephora website. This dataset consists of two types of metadata: one containing information about the products available on the website and the other containing product reviews.

TABLE I
PRODUCT INFO DATASET

product_id	product_name	brand_id	brand_name	category	rating
P473671	Fragrance Discovery Set	6342	19-69	Fragrance	3.6364

P473668	La Habana Eau de Parfum	6342	19-69	Fragrance	4.1538
P473662	Rainbow Bar Eau de Parfum	6342	19-69	Fragrance	4.2500
P473660	Kasbah Eau de Parfum	6342	19-69	Fragrance	4.4762
P473658	Purple Haze Eau de Parfum	6342	19-69	Fragrance	3.2308

TABLE II
Reviews Dataset

author_id	rating	product_id	product_name	review_text	review_title	is_recomended
1741593524	5	P504322	Gentle Hydra-Gel Face Cleanser	I use this with the Nudestix…	Taught me how to double cleanse!	1.0
31423088263	1	P420652	Lip Sleeping Mask Intense Hydration…	I bought this lip mask…	New Favorite Routine	0.0
5061282401	5	P420652	Lip Sleeping Mask Intense Hydration…	My review title says it all!…	Can't go wrong with any of them	1.0
6083038851	5	P420652	Lip Sleeping Mask Intense Hydration…	I've always loved this formula …	A must have !!!	1.0
47056667835	5	P420652	Lip Sleeping Mask Intense Hydration…	If you have dry cracked lips,…	Disappointed	1.0

Table I shows the contents of the dataset with information about all the products available on the Sephora website. Table II shows the contents of the product review dataset. The product metadata contains 2,420 rows and 28 columns. Meanwhile, the review metadata contains 1,094,411 rows and 19 columns.

B. Preprocessing

Preprocessing for recommendation systems using the matrix factorization method involves several stages. The first stage is filtering the product review dataset to ensure that only products categorized as skincare are included. This process is illustrated in Fig. 3.

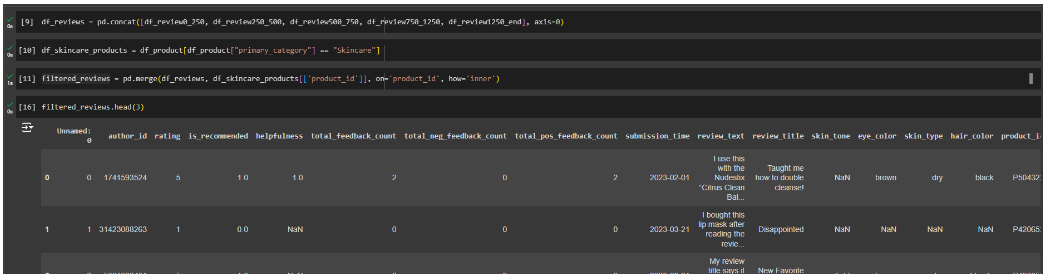

Fig 3. Data Separation Process

Once the dataset is filtered, the next step is to extract only the data required for the learning stage. Specifically, the dataset's columns used are limited to author_id, product_id, and rating, where the rating values range from 1 to 5.

The following step involves creating a user-item matrix, which is accomplished using the Reader class available in the Surprise library, as shown in Fig. 4. The user-item matrix serves as a representation of the relationships between users and items. Each cell in the matrix contains the rating or preference score given by a user to a specific item.

```
reader = Reader(rating_scale=(1, 5))
data = Dataset.load_from_df(filtered_reviews[["author_id",  "product_id", "rating"]], reader=reader)
```

Fig 4. Implementation of Reader Usage

An example of the preprocessing result is presented in Table III, which demonstrates the user-item matrix that reflects users' preferences for skincare products.

TABLE III
MATRIX USER-ITEM

	P504322	P420652	P420652	P420652	P420652
1741593524	5	0	0	0	0
31423088263	0	5	0	0	0
5061282401	0	0	0	0	0
6083038851	0	0	5	0	0
47056667835	0	0	0	0	5

C. Splitting Data

After the data preprocessing is complete, the User-Item matrix is divided into two parts: training data and test data. The training data is used to build the model, while the test data is used to evaluate the performance of the model. The proportions of the training data and test data are 80% and 20%, respectively.

D. Learning Matrix Factorization

In the learning context, the Surprise library (https://surpriselib.com) is used for implementing matrix factorization. This library provides various recommendation methods, including Non-negative Matrix Factorization (NMF) and Singular Value Decomposition (SVD). To achieve optimal model performance, hyperparameter tuning through cross-validation is carried out using GridSearchCV, which is also available in the Surprise library. The hyperparameter search involves several combinations of parameters, such as n_factors (searched values: 15, 50, and 100) and epochs (searched values: 10, 25, and 50). This tuning process aims to find the best configuration that maximizes the model's prediction accuracy, thereby providing more relevant recommendations to the users.

E. Performance Evaluation

After completing the learning phase with matrix factorization, the next step is to evaluate the model's performance using test data. The evaluation is carried out using several standard metrics, including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Precision, Recall, and F1 Score. Using these metrics is crucial for understanding how well the model can provide recommendations that are not only accurate but also relevant to user preferences.

IV. RESULTS AND DISCUSSION

A. Testing Scenario

The author conducted several testing scenarios to find the best results for this research. Hyperparameter tuning was used in these tests. The hyperparameter tuning utilized GridSearchCV on the complete dataset to find the optimal values for two parameters: factor and epoch, in order to achieve the best results for the learning process using both NMF and SVD algorithms.

The hyperparameter scenarios used the same combinations for learning tests with both NMF and SVD. The factor parameters tested were 15, 50, and 100, while the epoch parameters tested were 10, 25, and 50. To measure the best results from the combinations of factor and epoch, the author used the evaluation metrics RMSE and MAE. The results of this hyperparameter tuning can be seen in Table IV.

TABLE IV
Using Hyperparameter

Hyperparameter Tuning	NMF		SVD	
	MAE	RMSE	MAE	RMSE
'n_factors': 15, 'n_epochs': 10	0.8950	1.1714	0.8276	1.0833
'n_factors': 15, 'n_epochs': 25	0.7756	1.1152	0.7498	1.0215
'n_factors': 15, 'n_epochs': 50	0.8376	1.1335	0.7100	1.0089
'n_factors': 50, 'n_epochs': 10	0.7402	1.1996	0.8170	1.0725
'n_factors': 50, 'n_epochs': 25	1.5005	1.7255	0.7402	1.0164
'n_factors': 50, 'n_epochs': 50	0.6940	1.0965	0.7051	1.0011
'n_factors': 100, 'n_epochs': 10	0.7523	1.2224	0.8095	1.0657
'n_factors': 100, 'n_epochs': 25	1.7854	2.0553	0.7367	1.0148
'n_factors': 100, 'n_epochs': 50	0.7010	1.1193	0.7045	0.9986

In the learning process using the NMF method, the best results were achieved with the hyperparameter configuration of 50 factors and 50 epochs, yielding evaluation metrics of RMSE: 1.0965 and MAE: 0.6940. Meanwhile, for the SVD method, the best results were obtained with 100 factors and 50 epochs, resulting in RMSE: 0.9986 and MAE: 0.7045. These optimal hyperparameter configurations were used for subsequent stages of the research.

After determining the hyperparameters for the model, further testing was conducted based on the designed implementation framework. Precision and Recall were calculated for the generated recommendations, focusing on the top 10 recommendations for both methods with a threshold value of 3.5.

The calculated precision and recall values were satisfactory and demonstrated relevance for the users. The results of these evaluations are presented in Table V.

TABLE V
Precission and Recall Result

Metode	Precision	Recall	Treshold
NMF	0.7865	0.7987	3.5
SVD	0.8054	0.8144	3.5

B. Discussion

After identifying the optimal parameters for achieving maximum results, these parameters were applied to the learning process using data split into training and testing sets. The parameters used for the NMF method were 'n_factors': 50 and 'n_epochs': 50, while for the SVD method, the parameters were 'n_factors': 100 and 'n_epochs': 50. A threshold value of 3.5 was used for evaluating precision and recall. The results in Table VI show that the SVD method outperformed the NMF method.

TABLE VI
EVALUATION RESULT

Metode	MAE	RMSE	Precision	Recall	F-1 Score
NMF	0.7074	1.1052	0.7865	07987	0.7926
SVD	0.7190	1.0104	0.8054	0.8144	0.8099

The data in Table VI shows that the SVD method performed better than the NMF method. For NMF, the evaluation metrics for regression were MAE at approximately 0.7074 and RMSE at 1.1052. The classification metrics showed a precision of 0.7865, recall of 0.7987, and an F1 score of 0.7926.

In comparison, the SVD method achieved better results, with MAE at around 0.7190 and RMSE at 1.0104 for regression metrics. For classification metrics, precision was 0.8054, recall was 0.8144, and the F1 score was 0.8099.

Both methods demonstrated good evaluation results and effectively handled challenges such as the cold start problem and data sparsity in the dataset. However, the metrics indicate that the SVD method consistently outperformed NMF, making it the preferred choice in this study.

V. CONCLUSION

The skincare recommendation system using matrix factorization with Non-Negative Matrix Factorization (NMF) and Singular Value Decomposition (SVD) based on user reviews from the Sephora website was successfully implemented using the Surprise library. The testing results for both algorithms, NMF and SVD, show advantages in addressing the cold start problem and data sparsity present in the dataset. The recommendation system using the SVD algorithm demonstrated better performance with improved RMSE, Precision, Recall, and F1 Score compared to the NMF algorithm. However, the potential for further implementation of recommendation systems with matrix factorization is significant. Future research is expected to enhance the performance of recommendation systems by exploring more parameter comparisons and examining other algorithms such as SVD++, stochastic matrix factorization, and additional methods.

ACKNOWLEDGMENT

The authors would like to thank Telkom University for their support and the opportunity to conduct this research. We also extend our gratitude to all who have supported us throughout this study, including colleagues, mentors, and friends. Your invaluable assistance and encouragement have been crucial to the completion of this work.

REFERENCES

[1] K. Rodan, K. Fields, G. Majewski, and T. Falla, “Skincare Bootcamp: The Evolving Role of Skincare,” *Plast Reconstr Surg Glob Open*, vol. 4, no. 12S, p. e1152, Dec. 2016, doi: 10.1097/GOX.0000000000001152.

[2] Statista, “Skin Care - Indonesia,” https://www.statista.com/outlook/cmo/beauty-personal-care/skin-care/indonesia.

[3] J. Ben Schafer, J. Konstan, and J. Riedl, “Recommender systems in e-commerce,” in Proceedings of the 1st ACM conference on Electronic commerce, New York, NY, USA: ACM, Nov. 1999, pp. 158–166. doi: 10.1145/336992.337035.

[4] J. E. Prayogo, A. Suharso, and A. Rizal, “Analisis Perbandingan Model Matrix Factorization dan K-Nearest Neighbor dalam Mesin Rekomendasi Collaborative Berbasis Prediksi Rating,” Jurnal Informatika Universitas Pamulang, vol. 5, no. 4, p. 506, Dec. 2021, doi: 10.32493/informatika.v5i4.7379.

[5] M. H. Abdi, G. O. Okeyo, and R. W. Mwangi, “Matrix Factorization Techniques for Context-Aware Collaborative Filtering Recommender Systems: A Survey,” *Computer and Information Science*, vol. 11, no. 2, p. 1, Mar. 2018, doi: 10.5539/cis.v11n2p1.

[6] M. Ilhami and Suharjito, "Film recommendation systems using matrix factorization and collaborative filtering," in *2014 International Conference on Information Technology Systems and Innovation (ICITSI)*, IEEE, Nov. 2014, pp. 1–6. doi: 10.1109/ICITSI.2014.7048228.

[7] H. Ahmad Adyatma, Z. Baizal, and J. Telekomunikasi, "Book Recommender System Using Matrix Factorization with Alternating Least Square Method," *Journal of Information System Research*, vol. 4, no. 4, pp. 1286–1292, 2023, doi: 10.47065/josh.v4i4.3816.

[8] F. Nissa, A. H. Primandari, and A. K. Thalib, "COLLABORATIVE FILTERING APPROACH: SKINCARE PRODUCT RECOMMENDATION USING SINGULAR VALUE DECOMPOSITION (SVD)," *MEDIA STATISTIKA*, vol. 15, no. 2, pp. 139–150, Apr. 2023, doi: 10.14710/medstat.15.2.139-150.

[9] I. Yoshua, H. Bunyamin, and S. Si, "Pengimplementasian Sistem Rekomendasi Musik Dengan Metode Collaborative Filtering," 2021.

[10] G. Ferio, R. Intan, and S. Rostianingsih, "Sistem Rekomendasi Mata Kuliah Pilihan Menggunakan Metode User Based Collaborative Filtering Berbasis Algoritma Adjusted Cosine Similarity."

[11] A. A. Huda, R. Fajarudin, and A. Hadinegoro, "Sistem Rekomendasi Content-based Filtering Menggunakan TF-IDF Vector Similarity Untuk Rekomendasi Artikel Berita," *Building of Informatics, Technology and Science (BITS)*, vol. 4, no. 3, Dec. 2022, doi: 10.47065/bits.v4i3.2511.

[12] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowl Based Syst*, vol. 46, pp. 109–132, Jul. 2013, doi: 10.1016/j.knosys.2013.03.012.

[13] S. A. Gunarto, E. S. Honggara, and D. D. Purwanto, "Website Sistem Rekomendasi dengan Content Based Filtering pada Produk Perawatan Kulit," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 11, no. 3, p. 399, Jul. 2023, doi: 10.26418/justin.v11i3.59049.

[14] M. J. Pazzani, "Framework for collaborative, content-based and demographic filtering," *Artif Intell Rev*, vol. 13, no. 5, pp. 393–408, 1999, doi: 10.1023/A:1006544522159/METRICS.

[15] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, "An Algorithmic Framework for Performing Collaborative Filtering," *ACM SIGIR Forum*, vol. 51, no. 2, pp. 227–234, Aug. 2017, doi: 10.1145/3130348.3130372.

[16] S. Sharma, A. Sharma, Y. Sharma, and M. Bhatia, "Recommender system using hybrid approach," in *2016 International Conference on Computing, Communication and Automation (ICCCA)*, IEEE, Apr. 2016, pp. 219–223. doi: 10.1109/CCAA.2016.7813722.

[17] R. Pan *et al.*, "One-Class Collaborative Filtering," in *2008 Eighth IEEE International Conference on Data Mining*, IEEE, Dec. 2008, pp. 502–511. doi: 10.1109/ICDM.2008.16.

[18] P. B.Thorat, R. M. Goudar, and S. Barve, "Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System," *Int J Comput Appl*, vol. 110, no. 4, pp. 31–36, Jan. 2015, doi: 10.5120/19308-0760.

[19] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work - CSCW '94*, New York, New York, USA: ACM Press, 1994, pp. 175–186. doi: 10.1145/192844.192905.

[20] B. Santosa, "Use of Hybrid Methods in Making E-commerce Product Recommendation Systems to Overcome Cold Start Problems," *Telematika*, vol. 16, no. 1, Feb. 2023, doi: 10.35671/telematika.v16i1.2080.

[21] F. Wilhelm, "Matrix Factorization for Collaborative Filtering Is Just Solving an Adjoint Latent Dirichlet Allocation Model After All," in *Fifteenth ACM Conference on Recommender Systems*, New York, NY, USA: ACM, Sep. 2021, pp. 55–62. doi: 10.1145/3460231.3474266.

[22] B. Rocca, "Introduction to recommender systems," https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada.

[23] Xin Luo, Mengchu Zhou, Yunni Xia, and Qingsheng Zhu, "An Efficient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems," *IEEE Trans Industr Inform*, vol. 10, no. 2, pp. 1273–1284, May 2014, doi: 10.1109/TII.2014.2308433.

[24] Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," *Computer (Long Beach Calif)*, vol. 42, no. 8, pp. 30–37, Aug. 2009, doi: 10.1109/MC.2009.263.

[25] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., *Recommender Systems Handbook*. Boston, MA: Springer US, 2011. doi: 10.1007/978-0-387-85820-3.

[26] A. Lichterfeld, A. Hauss, C. Surber, T. Peters, U. Blume-Peytavi, and J. Kottner, "Evidence-Based Skin Care," *Journal of Wound, Ostomy & Continence Nursing*, vol. 42, no. 5, pp. 501–524, Sep. 2015, doi: 10.1097/WON.0000000000000162.

[27] A. Rahmawaty, "Peran Perawatan Kulit (Skincare) Yang Dapat Merawat Atau Merusak Skin Barrier," *Berkala Ilmiah Mahasiswa Farmasi Indonesia (BIMFI)*, vol. 7, no. 1, pp. 005–010, Sep. 2020, doi: 10.48177/bimfi.v7i1.32.

[28] L. Retno Hariatiningsih, "Penggunaan Skincare Dan Penerapan konsep Beauty 4.0 Pada Media Sosial (Studi Netnografi Wanita Pengguna Instagram)," *Journal Komunikasi*, vol. 11, no. 2, 2020, doi: 10.31294/jkom.

[29] D. Bokde, S. Girase, and D. Mukhopadhyay, "Matrix Factorization Model in Collaborative Filtering Algorithms: A Survey," *Procedia Comput Sci*, vol. 49, pp. 136–146, 2015, doi: 10.1016/j.procs.2015.04.237.