

Sistem Rekomendasi Skincare Menggunakan Matrix Factorization dengan Metode Non-Negative Matrix Factorization

1st Ahmad Indra Nurfauzi

Fakultas Informatika
Telkom University

Bandung, Indonesia

indranurfa@student.telkomuniversity.ac.id

2nd Agung Toto Wibowo

Fakultas Informatika
Telkom University

Bandung, Indonesia

agungtoto@telkomuniversity.ac.id

Abstrak — Perawatan kulit wajah sangat penting untuk menjaga kulit tetap bersih, sehat, dan bercahaya. Sistem rekomendasi, seperti Collaborative Filtering dan Content-Based Filtering, dapat memberikan saran produk skincare berdasarkan ulasan pengguna. Penelitian ini membandingkan dua model Matrix Factorization, yaitu Non-Negative Matrix Factorization (NMF), dan Singular Value Decomposition (SVD) untuk meningkatkan akurasi dan relevansi rekomendasi skincare. Hasil penelitian menunjukkan bahwa metode SVD lebih unggul, dengan Mean Absolute Error (MAE) sebesar 0,7190, Root Mean Square Error (RMSE) sebesar 1,0104, Precision sebesar 0,8054, Recall sebesar 0,8144, dan skor F-1 sebesar 0,8099. Sebagai perbandingan, metode NMF memiliki MAE sebesar 0,7074, RMSE sebesar 1,1052, Precision sebesar 0,7865, Recall sebesar 0,7987, dan skor F-1 sebesar 0,7926. Hasil pengujian menunjukkan bahwa kedua metode tersebut bisa menghasilkan rekomendasi sistem yang cukup akurat.

Kata kunci— Matrix Factorization, Collaborative Filtering, Skincare, Non-Negative Matrix Factorization, Singular Value Decomposition

I. PENDAHULUAN

Kulit wajah merupakan bagian kulit yang essensial untuk dirawat karena untuk menjaga tampilan agar terlihat bersih, sehat dan bersinar. Perawatan kulit wajah atau biasa disebut dengan istilah skincare merupakan suatu tindakan yang dapat dilakukan untuk mempertahankan kondisi kulit agar tetap bersih, sehat dan bersinar [1]. Skincare dilakukan dengan menggunakan berbagai macam produk-produk pendukung yang memiliki kandungan sesuai dengan kondisi kulit. Produk-produk tersebut nantinya akan diaplikasikan ke dalam rutinitas perawatan yang dapat menjaga agar kulit tetap sehat, bersih dan terawat.

Menurut data dari Portal data pasar dan konsumen internasional, Statista pada tahun 2023 [2], pasar industri kosmetik di Indonesia diperkirakan akan tumbuh sebesar 4,59 persen per tahun dalam periode 2023-2028. Proyeksi ini meliputi produk seperti perawatan kulit (skincare) dan juga perawatan diri (personal care). Selain itu, Badan Pengawas Obat dan Makanan (BPOM) menyebutkan bahwa jumlah industri kosmetik meningkat dari 819 unit usaha pada tahun

2021 menjadi 913 unit usaha pada tahun 2022, menunjukkan pertumbuhan sebesar 20,6 persen.

Sistem Rekomendasi adalah sebuah sistem yang mampu menyarankan item atau konten yang menarik, yang kemungkinan besar akan dipilih, digunakan, atau dibeli oleh pengguna [3]. Sistem rekomendasi digunakan oleh situs E-commerce seperti Shephora, Tokopedia, Shopee untuk memberikan saran produk kepada pelanggan mereka. Produk dapat direkomendasikan berdasarkan penjualan teratas secara keseluruhan di situs tersebut, berdasarkan demografi pelanggan, atau berdasarkan analisis perilaku belanja pelanggan di masa lalu sebagai prediksi untuk perilaku belanja di masa depan [4].

Dalam sistem rekomendasi, terdapat dua metode yang paling umum digunakan yakni Collaborative Filtering dan Content Based Filtering. Dalam metode collaborative filtering, rekomendasi dibuat berdasarkan perkiraan rating suatu item dari pengguna lain yang memiliki preferensi serupa [5]. Content Based Filtering adalah salah satu metode paling awal dan paling populer dalam sistem rekomendasi. Prinsip dari metode ini adalah merekomendasikan objek yang memiliki kesamaan dengan objek lain yang disukai oleh pengguna di masa lalu. Kesamaan antara objek ditentukan dari nilai-nilai karakteristik objek tersebut [6].

Keterkaitan antara review positif pada produk skincare dan kemudahan dalam direkomendasikan kepada pengguna menjadi semakin signifikan di tengah arus cepatnya produk skincare yang beredar saat ini. Dengan jumlah produk skincare yang terus berkembang, konsumen sering kali menghadapi kesulitan dalam memilih produk yang cocok dengan kebutuhan dan preferensi mereka. Dalam konteks ini, sistem rekomendasi collaborative filtering memegang peranan penting. Dengan memanfaatkan informasi review atau rating yang diberikan oleh pengguna terhadap produk, sistem ini dapat secara efektif menyaring dan merekomendasikan produk skincare yang berkualitas tinggi kepada pengguna. Kemampuan sistem collaborative filtering untuk menganalisis dan menggabungkan data historis rating dari berbagai pengguna memungkinkan pengguna mendapatkan rekomendasi yang personal dan relevan. Dengan demikian, sistem rekomendasi collaborative filtering memberikan solusi cerdas untuk membantu pengguna menavigasi dalam keragaman produk skincare yang ada,

memastikan bahwa produk dengan review positif dapat dengan mudah diakses dan dinikmati oleh konsumen.

Non-negative Matrix Factorization (NMF) merupakan salah satu metode pada Collaborative Filtering (CF). Non-negative Matrix Factorization (NMF) metode faktorisasi matriks yang dilakukan dengan cara dekomposisi sebuah matrix menjadi 2 buah matrix berukuran yang lebih kecil. NMF adalah metode faktorisasi matriks yang mempertahankan asumsi bahwa semua nilai dalam matriks faktor adalah non-negatif [7]. Dalam sistem rekomendasi, matriks ini berisi data seperti rating atau ulasan produk yang diberikan oleh pengguna. Selain metode NMF, Singular Value Decomposition (SVD) juga merupakan metode yang populer dalam sistem rekomendasi. SVD adalah faktorisasi matriks yang dilakukan dengan cara mendekomposisi sebuah matrix menjadi 3 buah matrix. SVD merupakan sebuah metode yang dikembangkan oleh Simon Funk ketika mengikuti kompetisi Netflix Prize pada tahun 2007 [8].

Sistem rekomendasi collaborative filtering berperan penting dalam membantu pengguna memilih produk skincare yang sesuai di tengah banyaknya pilihan yang tersedia. Dengan menganalisis review dan rating dari berbagai pengguna, sistem ini dapat memberikan rekomendasi yang personal dan relevan, memastikan produk berkualitas tinggi lebih mudah ditemukan oleh konsumen. Oleh karena itu, penelitian tentang sistem rekomendasi produk skincare berdasarkan rating pengguna pada web Sephora dengan menggunakan Matrix Factorization dengan membandingkan 2 model, yaitu Non-Negatif Matrix Factorization dan Singular Value Decomposition. Sistem rekomendasi skincare yang dapat memberikan rekomendasi yang lebih tepat dan sesuai dengan kebutuhan pengguna, meningkatkan pengalaman pengguna dalam mengambil keputusan memilih skincare yang cocok sesuai dengan produk yang direkomendasikan, serta meningkatkan efisiensi dan efektivitas dalam memberikan rekomendasi.

II. KAJIAN TEORI

Penelitian yang dilakukan oleh Adyatma dkk (2023) [9] menggunakan dataset dari Goodreads berupa data buku dan data penilaian. Dalam penelitian ini menggunakan metode Collaborative Filtering dengan membandingkan 2 metode yang berbeda yaitu Singular Value Decomposition (SVD) dan Alternating Least Square (ALS). Berdasarkan hasil penelitian, SVD berhasil mencapai nilai akurasi yang lebih baik dengan nilai RMSE sekitar 0,86822, nilai MAE sekitar 0,6903, F1-Score sekitar 0,827923, dan Precision sekitar 0,568347. Sementara itu, metode ALS mendapatkan nilai RMSE sekitar 1,09320, nilai MAE sekitar 0,86479, F1-Score sekitar 0,000304, dan Precision sekitar 0,000596.

Penelitian yang dilakukan oleh Nissa dkk (2023) [10] melakukan sebuah penelitian sistem rekomendasi skincare dengan menggunakan metode Collaborative Filtering dengan membandingkan 2 metode yang berbeda yaitu Singular Value Decomposition (SVD) dan Alternating Least Square (ALS). Dari penelitian ini didapatkan hasil bahwa metode ALS, ketika dibandingkan dengan metode SVD, menghasilkan nilai RMSE yang lebih tinggi dengan 10-fold CV, yaitu sebesar 1,00949 dibandingkan dengan 1,00915.

Penelitian yang dilakukan oleh Yoshua dkk (2021) [11] melakukan sebuah penelitian sistem rekomendasi musik

menggunakan metode Collaborative Filtering. Dalam penelitian ini dievaluasi menggunakan metrik Root Mean Squared Error (RMSE) dan Mean Absolute Error (MAE) untuk mengukur kinerja dari sistem rekomendasi yang dibuat. Hasil evaluasi menunjukkan bahwa metode Singular Value Decomposition (SVD)++ memberikan performa terbaik dengan nilai RMSE sebesar 0,0386 dan MAE sebesar 0,0228. SVD juga menunjukkan kinerja yang baik dengan nilai RMSE sebesar 0,0457 dan MAE sebesar 0,0291. Di sisi lain, metode Nearest Neighbors (KNN) dengan metode *User-User* dan *Item-Item* menggunakan Cosine Similarity atau Pearson Similarity menunjukkan hasil yang seragam dengan nilai RMSE dan MAE sebesar 0,1295 dan 0,0526. Metode Probabilistic Matrix Factorization (PMF) dan Non-Negative Matrix Factorization (NMF) menunjukkan nilai RMSE dan MAE sebesar 0,1090 dan 0,0325, serta 0,0677 dan 0,0289 secara berturut-turut.

A. Sistem Rekomendasi

Sistem rekomendasi adalah sistem yang dapat memberikan informasi dan rekomendasi untuk membantu pengguna dalam membuat keputusan berdasarkan data yang telah ada sebelumnya [12]. Tujuan dari sistem rekomendasi adalah untuk memberikan konten (item) yang efektif dan berarti kepada pengguna yang aktif didalam platform [13]. Sistem rekomendasi dapat menggunakan berbagai metode [14]. Metode-metode tersebut digunakan untuk merekomendasikan produk, layanan, atau informasi kepada pengguna berdasarkan preferensi mereka. Metode tersebut meliputi collaborative filtering, demographic filtering, content-based filtering, dan hybrid filtering.

Content-based filtering memberikan rekomendasi berdasarkan pilihan pengguna yang dibuat di masa lalu [15]. Content-based filtering juga menghasilkan rekomendasi dengan menggunakan konten dari objek yang dimaksudkan untuk direkomendasikan. Demographic filtering berdasarkan prinsip bahwa individu dengan atribut pribadi tertentu yang sama (jenis kelamin, usia, negara, dll.) juga akan memiliki preferensi yang sama [16]. Collaborative Filtering memungkinkan pengguna memberikan peringkat tentang serangkaian elemen (misalnya, video, lagu, film, dll. di situs web berbasis CF) sehingga ketika informasi yang cukup disimpan dalam sistem, kita dapat memberikan rekomendasi kepada setiap pengguna berdasarkan informasi yang diberikan oleh pengguna yang dianggap memiliki kesamaan paling banyak dengan mereka [17]. Sementara itu Hybrid Filtering merupakan kombinasi dari collaborative filtering dengan demographic filtering atau collaborative filtering dengan content-based filtering [18].

B. Collaborative Filtering

Collaborative Filtering adalah metode rekomendasi yang populer yang menghasilkan prediksi dan rekomendasi berdasarkan peringkat atau opini dari pengguna lain di dalam sistem [19]. Metode Collaborative Filtering menggunakan data yang didasarkan pada kesamaan karakteristik konsumen untuk memberikan informasi berdasarkan pola dari kelompok konsumen yang mirip, sehingga memungkinkan pemberian informasi baru kepada konsumen [20].

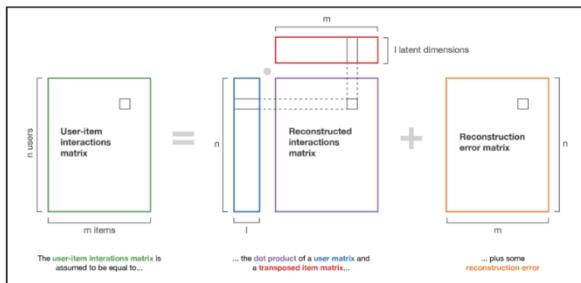
Sistem Collaborative Filtering, contohnya seperti GroupLens [21], memanfaatkan data peringkat pengguna untuk mengalkulasikan kesamaan atau bobot antara

pengguna atau item. Berdasarkan nilai kesamaan yang dihitung tersebut, sistem ini membuat prediksi atau rekomendasi. Metode Collaborative Filtering berbasis memori, yang sering disebut sebagai CF, khususnya diintegrasikan ke dalam platform komersial [22]. Contoh yang populer dari CF berbasis memori adalah penerapan Collaborative Filtering untuk memberikan rekomendasi terkait item-item.

Teknik CF berbasis model, di sisi lain, menggunakan data peringkat mentah untuk memperkirakan atau mempelajari model untuk membuat prediksi [11]. Model ini dapat berupa algoritma data mining atau machine learning. Teknik CF berbasis model terkenal melibatkan Bayesian belief nets (BNs) CF models, clustering CF models, dan latent semantic CF models [23].

C. Matrix Factorization

Matrix factorization beroperasi dengan cara melakukan dekomposisi matriks berukuran besar menjadi matriks yang lebih kecil. Hasil akhir dari matriks ini merupakan hasil dari perkalian titik (dot product) antara matriks pengguna dan transpose dari matriks item [24].



GAMBAR 1

Ilustrasi matrix factorization

1. Non-negative Matrix Factorization (NMF)

Non-negative Matrix Factorization adalah metode faktorisasi matriks yang dilakukan dengan cara dekomposisi sebuah matrix menjadi 2 buah matrix berukuran yang lebih kecil. NMF adalah metode faktorisasi matriks yang mempertahankan asumsi bahwa semua nilai dalam matriks faktor adalah non-negatif. Pendekatan ini berguna terutama ketika data bersifat non-negatif, seperti dalam kasus peringkat pengguna terhadap item. Rumus prediksi NMF memiliki rumus (1):

$$\hat{r}_{ui} = q_i^T p_u \quad (1)$$

Prosedur optimisasi adalah stochastic gradient descent (dengan regularisasi) dengan pemilihan ukuran langkah tertentu yang memastikan faktor-faktor tetap tidak negatif, asalkan nilai awalnya juga positif. Rumus perubahan nilai menggunakan model regularized single-element-based NMF (RSNMF) yang didasarkan pada temuan Xin Luo dkk (2014) [25]. Pada setiap langkah dari prosedur SGD, faktor untuk pengguna dan item diperbarui yang memiliki rumus (2):

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

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

2. Singular Value Decomposition (SVD)

SVD adalah salah satu metode matrix factorization yang populer di sistem rekomendasi. SVD merupakan sebuah metode yang dikembangkan oleh Simon Funk ketika mengikuti kompetisi Netflix Prize pada tahun 2007 [26]. SVD adalah faktorisasi matriks yang dilakukan dengan cara mendekomposisi sebuah matrix menjadi 3 buah matrix. Konsep dekomposisi matriks dari matrix factorization diterapkan oleh metode SVD dalam bentuk rumus untuk menghasilkan nilai prediksi yang berdasarkan pada Ricci dkk (2011) [27]. Bentuk dari rumus prediksi (3):

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

Untuk memperkirakan atau menghitung semua nilai rating yang belum diketahui, SVD dapat melakukannya dengan meminimalkan regularized square error menggunakan rumus yang dijelaskan dalam Rumus (4):

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

Untuk mengurangi kesalahan yang dihasilkan oleh model algoritma dalam membuat prediksi, diperlukan langkah-langkah yang dilakukan oleh gradient descent menggunakan rumus-rumus yang dijelaskan dalam Rumus (5):

$$\begin{aligned} b_u &\leftarrow b_u + \gamma(e_{ui} - \lambda b_u) \\ b_i &\leftarrow b_i + \gamma(e_{ui} - \lambda b_i) \\ p_u &\leftarrow p_u + \gamma(e_{ui} \cdot q_i - \lambda p_u) \\ q_i &\leftarrow q_i + \gamma(e_{ui} \cdot p_u - \lambda q_i) \end{aligned} \quad (5)$$

D. Skincare

Secara umum, skincare adalah aktivitas merawat kulit bagian luar tubuh dengan menggunakan produk-produk tertentu untuk menjaga kesehatan dan penampilan kulit [28]. Skincare tidak hanya pada kalangan Wanita saja tetapi juga digunakan pada kalangan Pria. Produk skincare memiliki berbagai jenis, namun yang paling umum meliputi pembersih, pencuci wajah, toner, krim pelembab, serum, dan tabir surya [29]. Facial wash berfungsi sebagai pembersih wajah dan tersedia dalam berbagai variasi untuk berbagai macam kondisi jenis kulit wajah. Toner digunakan untuk menyegarkan kulit dan mengurangi minyak berlebih di wajah. Krim pelembab sangat penting untuk menjaga kelembaban kulit wajah serta mencegah kerusakan akibat pemakaian make-up. Sunscreen juga diperlukan untuk melindungi kulit dari efek berbahaya sinar UV yang berasal dari matahari [29].

E. Metrik Evaluasi

1. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) digunakan untuk mengukur perbedaan sebagai nilai rata-rata antara prediksi algoritma dan peringkat sebenarnya yang akan diberikan oleh pengguna [30]. MAE didapatkan dengan rumus (6):

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

2. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) adalah metrik yang digunakan untuk mengukur akurasi model prediksi dengan menghitung akar dari rata-rata kuadrat dari

semua kesalahan (selisih antara nilai prediksi dan nilai sebenarnya) [30]. RMSE didapatkan dengan rumus (7):

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

3. Precision

Precision adalah perbandingan antara total jumlah item relevan dengan jumlah item yang direkomendasikan [30]. Precision dapat dihitung dengan rumus (8):

$$Precision = \frac{|\text{Interesting Items} \cap \text{Recommended Items}|}{|\text{Recommended Items}|} \quad (8)$$

4. Recall

Recall adalah perbandingan antara total jumlah *item* relevan yang telah direkomendasikan dengan total keseluruhan *item* relevan [30]. Recall dihitung menggunakan rumus (9):

$$Recall = \frac{|\text{Interesting Items} \cap \text{Recommended Items}|}{|\text{Interesting Items}|} \quad (9)$$

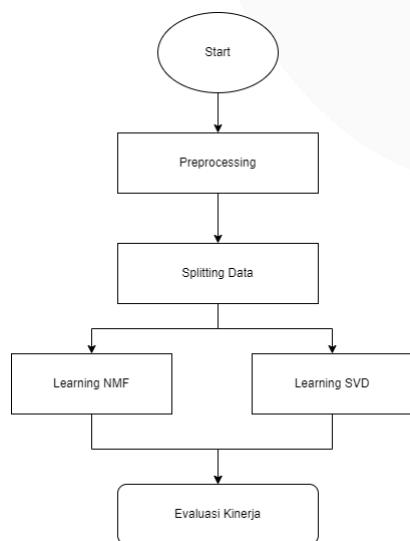
5. F-1 Score

F-1 Score adalah kombinasi dari precision dan recall, digunakan untuk mengukur keseimbangan antara precision dan recall dan sangat berguna ketika ada distribusi kelas yang tidak rata [30]. F-1 Score didapatkan dengan rumus (10):

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (10)$$

III. METODE

Dalam perancangan sistem rekomendasi skincare menggunakan matrix factorization dengan menggunakan perhitungan NMF dan SVD. Terdapat bagian penting seperti preprocessing, splitting data dan juga perhitungan menggunakan metode NMF dan SVD. Dalam merancang arsitektur suatu sistem, diperlukan flowchart yang menunjukkan langkah-langkah bagaimana sistem berjalan. Pada Gambar 2 merupakan flowchart rancangan sistem yang dibangun:



GAMBAR 2

Flowchart Desain Sistem

A. Dataset

Dalam penelitian ini, dataset yang digunakan berasal dari web Kaggle, data tersebut diambil dari situs web Sephora menggunakan Python scraper pada bulan maret 2023 [31]. Dataset ini terdiri dari dua jenis dataset, yaitu dataset yang mengandung informasi produk yang tersedia di situs web dan dataset yang berisi review produk.

TABEL 1
Dataset Info Produk

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	...
...

TABEL 2
Dataset Rating

author_id	rating	product_id	product_name	...
1741593524	5	P504322	Gentle Hydra-Gel Face Cleanser	...
31423088263	1	P420652	Lip Sleeping Mask Intense Hydration...	...
5061282401	5	P420652	Lip Sleeping Mask Intense Hydration...	...
6083038851	5	P420652	Lip Sleeping Mask Intense Hydration...	...
47056667835	5	P420652	Lip Sleeping Mask Intense Hydration...	...
...

Tabel 1 menunjukkan isi dari dataset informasi dari semua produk kecantikan yang tersedia di web Sephora, yang terdiri dari 2.420 baris dan 28 kolom, termasuk nama produk dan merek, harga, bahan-bahan, rata-rata rating produk, serta semua informasi produk lainnya. Tabel 2 menunjukkan isi dari dataset review produk, yang berisi 1.094.411 baris dan 19 kolom, mencakup review pengguna dari semua produk dalam kategori Skincare, termasuk penampilan pengguna dan penilaian review oleh pengguna lain.

B. Preprocessing

Preprocessing untuk sistem rekomendasi dengan metode matrix factorization ada beberapa tahap yang dilakukan. Tahapan pertama yaitu melakukan penyaringan data di dataset review produk untuk memastikan produk yang ada di dataset tersebut hanyalah yang berkategori Skincare, yang dapat dilihat pada GAMBAR 3.

```
[1]: reviews = pd.concat([recalled_ids, df_reviewed_ids, df_reviews_ids, df_myseen_ids, df_recalled_ids], axis=0)

[2]: df_myseen_products = product[(product['category'].str.contains('electronics')) | (product['category'].str.contains('books'))]

[3]: filtered_reviews = reviews[~(df_myseen_products['product_id'].isin(reviews['product_id'])) & ~reviews['user_id'].isin(df_myseen_products['user_id'])]

[4]: filtered_reviews.head(5)

[5]: filtered_reviews.info()

[6]: filtered_reviews.describe()

[7]: filtered_reviews[['rating', 'is_recommended', 'helpfulness', 'total_feedback_count', 'total_avg_feedback_count', 'total_pos_feedback_count', 'submittime', 'time_rev_rec', 'rec_weight', 'skip_time', 'skip_rate', 'skip_type', 'product_id']].head(10)

[8]: filtered_reviews.info()

[9]: filtered_reviews.describe()

[10]: filtered_reviews[['rating', 'is_recommended', 'helpfulness', 'total_feedback_count', 'total_avg_feedback_count', 'total_pos_feedback_count', 'submittime', 'time_rev_rec', 'rec_weight', 'skip_time', 'skip_rate', 'skip_type', 'product_id']].head(10)
```

GAMBAR 3

Proses Pemisahan Data

Lalu setelah tahapan tersebut dilakukan, selanjutnya memisahkan data yang hanya digunakan pada tahapan learning yakni pada dataset review kolom yang digunakan hanya author_id, product_id dan rating, yang dimana rating bernilai angka 1-5. Selanjutnya pembuatan matriks user-item, kedua tahapan tersebut dilakukan dengan menggunakan Reader yang terdapat pada library Surprise, seperti pada Gambar 4.

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

GAMBAR 4

Implementasi Penggunaan Reader

Matriks user-item berfungsi sebagai representasi hubungan antara pengguna dan item, di mana setiap sel matriks menyimpan nilai peringkat atau preferensi pengguna terhadap item tertentu. Pada table 3 merupakan contoh hasil dari preprocessing yang telah menjadi matriks User-Item yang mencerminkan preferensi pengguna terhadap produk skincare.

TABEL 3
Matrix User-Item

	P504322	P420652	P420652
1741593524	5	0	0
31423088263	0	5	0
5061282401	0	0	0

C. Splitting Data

Setelah proses preprocessing data selesai, matriks User-Item dibagi menjadi dua bagian: data training dan data test. Data training digunakan untuk membangun model, sedangkan data test digunakan untuk mengevaluasi kinerja model tersebut. Proporsi data training dan data test masing-masing adalah 80% dan 20%.

D. Learning Matrix Factorization

Dalam konteks learning, penggunaan library Surprise (<https://surpriselib.com>) digunakan untuk implementasi matrix factorization. Library ini menyediakan berbagai metode rekomendasi, termasuk Non-negative Matrix Factorization (NMF) dan Singular Value Decomposition (SVD). Untuk mendapatkan performa model yang optimal, dilakukan hyperparameter tuning cross validation dengan menggunakan GridSearchCV yang juga tersedia pada library Surprise. Pencarian hyperparameter dilakukan dengan beberapa kombinasi parameter, seperti n_factors yang dicari yakni 15, 50 dan 100 dan parameter epoch yang dicari adalah

10, 25 dan 50. Proses tuning ini bertujuan untuk menemukan konfigurasi terbaik yang dapat memaksimalkan akurasi prediksi model, sehingga dapat memberikan rekomendasi yang lebih relevan bagi pengguna.

E. Evaluasi

Setelah melalui tahap learning dengan matrix factorization, langkah selanjutnya adalah mengevaluasi kinerja model menggunakan data uji. Evaluasi dilakukan dengan memanfaatkan beberapa metrik standar, antara lain: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Precision, Recall, dan F-1 Score. Evaluasi menggunakan metrik-metrik ini memiliki peran penting dalam memberikan pemahaman tentang sejauh mana model dapat memberikan rekomendasi yang tidak hanya akurat tetapi juga relevan dengan preferensi pengguna.

IV. HASIL DAN PEMBAHASAN

Penulis melakukan beberapa skenario pengujian yang digunakan untuk mencari hasil terbaik terhadap pengujian penelitian ini, penulis menggunakan hyperparameter tuning di dalam pengujian ini. Penggunaan tuning hyperparameter disini menggunakan GridSearchCV pada data utuh yang digunakan untuk mencari dua parameter yang berbeda yakni factor dan epoch untuk mencari hasil yang terbaik untuk proses learning menggunakan metode NMF dan juga SVD.

Skenario penggunaan hyperparameter disini menggunakan kombinasi yang sama untuk pengujian learning menggunakan NMF maupun SVD. Untuk parameter factor yang dicari adalah 15, 50 dan 100. Sedangkan untuk parameter epoch yang dicari adalah 10, 25 dan 50. Dan juga untuk mengukur hasil yang terbaik dari kombinasi antara factor dan epoch tersebut, penulis menggunakan hasil evaluasi metric RMSE dan MAE. Hasil dari hyperparameter tuning tersebut dapat dilihat pada tabel 4.

TABEL 4
Hasil 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

Pada learning menggunakan metode NMF menemukan hasil yang terbaik adalah penggunaan hyperparameter factor 50 dan epoch 50 dengan hasil evaluasi yang didapat RMSE 1,0965 dan MAE 0,6940, sedangkan pada learning menggunakan metode SVD menemukan hasil yang terbaik adalah penggunaan hyperparameter factor 100 dan epoch 50 dengan hasil evaluasi yang didapat RMSE 0,9986 dan MAE 0,7045. Hasil yang didapat daripada pengujian hyperparameter tersebut digunakan untuk kelanjutan penelitian yang dilakukan.

Setelah mengetahui parameter yang digunakan agar hasil maksimal, lalu parameter tersebut digunakan untuk learning pada data yang sudah di pisahkan menjadi data training dan data test. Parameter yang digunakan pada metode NMF yakni 'n_factors': 50, 'n_epochs': 50 dan pada metode SVD yakni 'n_factors': 100, 'n_epochs': 50. Dan threshold yang digunakan untuk mengukur evaluasi precision dan recall dengan angka 3,5.

TABEL 5
Hasil Learning NMF dan SVD

Metode	MAE	RMSE	Precision	Recall	F-1 Score
NMF	0,7074	1,1052	0,7865	0,7987	0,7926
SVD	0,7190	1,0104	0,8054	0,8144	0,8099

Pada data pada tabel 5, memperlihatkan bahwa metode SVD lebih unggul dibandingkan dengan metode NMF. Untuk metode NMF memiliki nilai evaluasi metric regressi nilai MAE sekitar 0,7074 dan untuk nilai RMSE sekitar 1,1052, sedangkan untuk nilai evaluasi metric classifikasi Precision yang diperoleh sekitar 0,7865, untuk nilai Recall yang diperoleh sekitar 0,7987 dan untuk nilai F-1 nilai yang diperoleh sekitar 0,7926.

Untuk metode SVD memiliki nilai evaluasi metric regressi nilai MAE sekitar 0,7190 dan untuk nilai RMSE sekitar 1,0104, sedangkan untuk nilai evaluasi metric classifikasi Precision yang diperoleh sekitar 0,8054, untuk nilai Recall yang diperoleh sekitar 0,8144 dan untuk nilai F-1 nilai yang diperoleh sekitar 0,8099.

Berdasarkan hasil yang dipaparkan pada table 5 di atas baik metode NMF dan SVD memiliki hasil evaluasi yang cukup baik. Dapat diperlihatkan bahwasanya kedua metode tersebut memiliki kelebihan bisa mengatasi cold start problem dan data sparsity yang terjadi pada dataset yang digunakan. Dari berdasarkan data yang didapat dari hasil perhitungan evaluasi metric tersebut bisa terlihat bahwasanya metode SVD cenderung dominan lebih baik jika dibandingkan dengan NMF.

V. KESIMPULAN

Berdasarkan hasil penelitian, sistem rekomendasi skincare berbasis matrix factorization telah berhasil diimplementasikan menggunakan Library Surprise dengan metode Non-Negative Matrix Factorization (NMF) dan Singular Value Decomposition (SVD) menggunakan data ulasan pengguna dari situs Sephora. Metode SVD menunjukkan kinerja yang lebih baik dibandingkan NMF, terutama dalam hal nilai RMSE, Precision, Recall, dan F-1 Score. SVD mencatat nilai RMSE yang lebih rendah yaitu 1,0104 dibandingkan dengan NMF yang memiliki nilai RMSE sebesar 1,1052. Selain itu, SVD unggul dalam nilai Precision dengan skor 0,8054, lebih tinggi daripada NMF yang memiliki nilai Precision 0,7865. Dalam hal Recall, SVD juga lebih unggul dengan nilai 0,8144, sementara NMF hanya

mencapai 0,7987. Terakhir, F-1 Score SVD lebih tinggi dengan nilai 0,8099 dibandingkan dengan NMF yang memiliki nilai 0,7926. Penelitian lebih lanjut diharapkan dapat meningkatkan performa sistem rekomendasi ini dengan melakukan lebih banyak perbandingan parameter yang digunakan, serta dengan membandingkan metode SVD dan NMF dengan metode lain seperti SVD++, stochastic matrix factorization, atau metode lainnya untuk mencapai hasil yang lebih optimal.

REFERENSI

- [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. 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.
- [4] 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.
- [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] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999, doi: 10.1038/44565.
- [8] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, "Recommender Systems Handbook."
- [9] 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.
- [10] 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.
- [11] I. Joshua, H. Bunyamin, and S. Si, "Pengimplementasian Sistem Rekomendasi Musik Dengan Metode Collaborative Filtering," 2021.
- [12] G. Ferio, R. Intan, and S. Rostianingsih, "Sistem Rekomendasi Mata Kuliah Pilihan Menggunakan Metode User Based Collaborative Filtering Berbasis Algoritma Adjusted Cosine Similarity."
- [13] 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.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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/ICCA.2016.7813722.
- [19] 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.
- [20] 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.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] B. Rocca, "Introduction to recommender systems," <https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada>.
- [25] 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.
- [26] 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.
- [27] 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.
- [28] A. Lichtenfeld, 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.
- [29] 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.
- [30] 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.
- [31] N. Inky, "Sephora Products and Skincare Reviews." Accessed: Nov. 23, 2023. [Online]. Available: <https://www.kaggle.com/datasets/nadyinky/sephora-products-and-skincare-reviews>