

BAB 2 LANDASAN TEORI

2.1 Skincare

Skincare mengacu pada serangkaian produk dan teknik yang digunakan untuk menjaga kesehatan kulit, memperbaiki masalah kulit, dan melindungi kulit dari faktor lingkungan seperti polusi dan sinar UV [8]. Produk skincare biasanya dikategorikan berdasarkan fungsinya. Pembersih (*cleansers*) berfungsi untuk menghilangkan kotoran, minyak, dan sisa makeup dari kulit, seperti *facial wash* dan *micellar water*. Pelembap (*moisturizers*) membantu menjaga kelembapan kulit dan mencegah dehidrasi, misalnya krim pelembap atau *hydrating serum*. Eksfoliator (*exfoliators*) berguna untuk mengangkat sel kulit mati guna merangsang regenerasi kulit, termasuk di antaranya *scrub* dan *chemical exfoliants* seperti AHA/BHA.

Tabir surya (*sunscreen*) melindungi kulit dari paparan sinar UV yang dapat menyebabkan penuaan dini hingga kanker kulit; contohnya adalah *broad-spectrum SPF 50*. Selain itu, terdapat produk khusus (*treatment products*) yang dirancang untuk menangani permasalahan kulit tertentu, seperti jerawat, hiperpigmentasi, atau tanda-tanda penuaan, misalnya *retinol serum* dan *vitamin C cream*. Dalam sistem rekomendasi produk skincare, faktor-faktor seperti jenis kulit (kering, berminyak, kombinasi, sensitif) serta kandungan bahan aktif menjadi sangat krusial dalam menghasilkan rekomendasi yang relevan dan efektif [3].

2.2 Sistem Rekomendasi Produk Skincare

Sistem rekomendasi adalah sistem berbasis kecerdasan buatan yang dirancang untuk membantu pengguna menemukan item atau informasi yang relevan berdasarkan preferensi, pola interaksi, atau karakteristik pengguna [9]. Teknologi ini banyak digunakan dalam berbagai bidang, seperti e-commerce, layanan streaming, dan media sosial, termasuk dalam pemilihan produk perawatan kulit (skincare), untuk meningkatkan pengalaman pengguna dan mendukung pengambilan keputusan [10].

Dalam konteks produk skincare, sistem rekomendasi bertujuan untuk menyarankan produk yang sesuai dengan kebutuhan dan karakteristik kulit pengguna berdasarkan data historis, preferensi pengguna lain, dan/atau kandungan produk.

2.3 Content-Based Filtering (TF-IDF)

Content-Based Filtering (CBF) bekerja dengan membandingkan karakteristik item berdasarkan preferensi pengguna [11]. Salah satu teknik umum dalam CBF adalah **Term Frequency-Inverse Document Frequency** (TF-IDF), yang digunakan untuk menilai pentingnya suatu kata dalam suatu dokumen relatif terhadap keseluruhan kumpulan dokumen [12].

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t) \quad (2.1)$$

dengan:

$$\text{IDF}(t) = \log \left(\frac{N}{df(t)} \right) \quad (2.2)$$

Teknik ini digunakan dalam berbagai aplikasi, termasuk analisis teks dan rekomendasi berbasis deskripsi produk, seperti dalam pencocokan bahan aktif produk skincare [3].

2.4 Collaborative Filtering

Collaborative Filtering (CF) adalah metode dalam sistem rekomendasi yang menyarankan item kepada pengguna berdasarkan preferensi pengguna lain yang memiliki pola interaksi serupa [13]. CF bekerja dengan asumsi bahwa pengguna yang memiliki kesukaan terhadap item yang sama di masa lalu kemungkinan besar akan memiliki kesukaan yang serupa di masa depan.

Secara umum, CF terbagi menjadi dua pendekatan utama:

1. Memory-Based Collaborative Filtering

Mengukur kesamaan antar pengguna (*User-Based*) atau antar item (*Item-Based*) secara langsung dari data interaksi menggunakan teknik seperti *cosine similarity* atau *Pearson correlation* [14].

2. Model-Based Collaborative Filtering

Membangun model prediktif dari data interaksi menggunakan teknik machine learning, seperti Matrix Factorization dan Truncated Singular Value Decomposition (SVD), untuk menangkap representasi laten pengguna dan item. Truncated SVD mereduksi dimensi matriks interaksi user-item sehingga pola laten dapat dikenali dan kemiripan antar pengguna dihitung di ruang faktor laten menggunakan dot product [15].

Pendekatan model-based ini efektif mengatasi masalah sparsitas data dan meningkatkan efisiensi komputasi dibandingkan memory-based tradisional. Oleh karena itu, pada penelitian ini digunakan model-based CF dengan Truncated SVD untuk menghasilkan rekomendasi yang personal dan relevan [16].

2.5 Weighted Hybrid Similarity

Hybrid Filtering mengombinasikan **CBF** dan **CF** untuk meningkatkan akurasi rekomendasi [17]. Dalam metode **Weighted Hybrid Similarity**, skor kesamaan dari kedua pendekatan digabungkan menggunakan bobot adaptif:

$$S_{\text{Hybrid}}(A, B) = \alpha S_{\text{CBF}}(A, B) + (1 - \alpha) S_{\text{CF}}(A, B) \quad (2.3)$$

dengan α sebagai bobot yang dapat disesuaikan untuk menyeimbangkan kedua metode. Pendekatan ini memungkinkan sistem untuk lebih fleksibel dalam menangkap karakteristik individu maupun pola perilaku kolektif [18].

2.6 Sentiment Analysis (VADER)

Sentiment Analysis merupakan proses analisis opini dalam teks, dengan tujuan mengidentifikasi dan mengklasifikasikan sentimen menjadi tiga kategori utama, yaitu positif, negatif, dan netral [19]. Pada penelitian ini, analisis sentimen diterapkan terhadap ulasan pengguna produk skincare sebagai salah satu elemen penting dalam sistem rekomendasi berbasis hybrid.

Metode analisis sentimen yang digunakan adalah VADER (Valence Aware Dictionary and sEntiment Reasoner), yang diimplementasikan melalui pustaka *Sentiment Intensity Analyzer* pada *Natural Language Toolkit* (NLTK) [20]. VADER merupakan leksikon berbasis aturan yang dirancang khusus untuk teks sosial dan ulasan dalam bahasa Inggris. Keunggulan utama VADER adalah kemampuannya dalam menangani ekspresi emosional informal, seperti emoji, penggunaan huruf kapital, intensifikasi (booster words), dan tanda baca (seperti "!!!" atau "?!").

VADER menghasilkan empat skor sentimen: *positive*, *negative*, *neutral*, dan *compound*. Skor *compound* merupakan hasil normalisasi dari total skor sentimen dan berada pada rentang -1 (sangat negatif) hingga +1 (sangat positif). Skor ini digunakan untuk menentukan polaritas keseluruhan dari sebuah teks. Pengklasifikasian sentimen berdasarkan nilai *compound* pada VADER mengacu pada aturan berikut:

- Positif: $compound \geq 0.05$
- Netral: $-0.05 < compound < 0.05$
- Negatif: $compound \leq -0.05$

Batas threshold ini cukup umum digunakan sesuai dengan dokumentasi resmi pengembang VADER serta berbagai publikasi dan literatur penelitian yang memanfaatkan VADER untuk analisis sentimen. Penjelasan nilai threshold ini juga termuat secara eksplisit pada repositori resmi VADER dan publikasi orisinal. [21, 22]

Penggunaan VADER dipilih karena efisiensi dan akurasi tinggi dalam menangani teks informal serta kemudahannya diintegrasikan dalam pipeline pemrosesan teks. Metode ini juga telah digunakan secara luas dalam penelitian-penelitian serupa terkait sistem rekomendasi berbasis ulasan pengguna dan terbukti efektif dalam meningkatkan kualitas prediksi. [20].

2.7 Ranking Metrics

Precision dan **Recall** merupakan dua metrik utama yang digunakan untuk mengevaluasi performa model klasifikasi, termasuk pada sistem klasifikasi sentimen [23].

- **Precision** mengukur proporsi prediksi positif yang benar-benar relevan (benar). Dengan kata lain, precision menunjukkan seberapa banyak hasil positif yang dihasilkan model adalah benar. Rumus precision adalah:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2.4)$$

di mana:

- *TP (True Positive)*: jumlah data positif yang diprediksi benar oleh model,
- *FP (False Positive)*: jumlah data negatif yang salah diprediksi sebagai positif oleh model.

- **Recall** mengukur proporsi data positif yang berhasil ditemukan oleh model dari seluruh data yang sebenarnya positif. Recall juga dikenal sebagai

sensitivitas. Rumus recall adalah:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.5)$$

di mana:

- *TP (True Positive)*: jumlah data positif yang diprediksi benar oleh model,
- *FN (False Negative)*: jumlah data positif yang gagal diprediksi oleh model.

Precision yang tinggi menunjukkan bahwa prediksi positif model sangat tepat, sedangkan recall yang tinggi menunjukkan bahwa model mampu menemukan sebagian besar data positif yang ada. Kedua metrik ini penting untuk menilai kinerja model klasifikasi, terutama jika data tidak seimbang atau konsekuensi kesalahan prediksi berbeda pada masing-masing kelas [24].

- **Hit Rate** adalah metrik yang digunakan untuk mengukur seberapa sering sistem rekomendasi berhasil memberikan setidaknya satu item relevan dalam daftar rekomendasi kepada pengguna. Hit Rate pada top-*K* (*Hit Rate@K*) didefinisikan sebagai persentase pengguna yang menerima minimal satu item relevan di antara *K* hasil teratas [25]. Rumus Hit Rate@*K* adalah:

$$\text{Hit Rate@K} = \frac{\text{Jumlah pengguna dengan minimal satu item relevan di top-}K}{\text{Total pengguna}} \quad (2.6)$$

Nilai Hit Rate yang tinggi menunjukkan bahwa sistem mampu memberikan rekomendasi yang bermanfaat bagi sebagian besar pengguna.

- **Normalized Discounted Cumulative Gain (NDCG)** adalah metrik yang digunakan untuk mengevaluasi kualitas urutan hasil rekomendasi dengan mempertimbangkan posisi item relevan dalam daftar. NDCG memberikan bobot lebih tinggi pada item relevan yang muncul di posisi atas [25]. Rumus NDCG pada top-*K* (*NDCG@K*) adalah:

$$\text{DCG@K} = \sum_{i=1}^K \frac{2^{rel_i} - 1}{\log_2(i + 1)} \quad (2.7)$$

$$\text{NDCG@K} = \frac{\text{DCG@K}}{\text{IDCG@K}} \quad (2.8)$$

di mana rel_i adalah relevansi item pada posisi ke- i , dan $IDCG@K$ adalah nilai DCG ideal jika semua item relevan berada di posisi teratas. NDCG bernilai antara 0 dan 1, dengan nilai lebih tinggi menunjukkan urutan rekomendasi yang lebih optimal.

- **Mean Average Precision (MAP)** adalah metrik yang digunakan untuk mengevaluasi kualitas sistem perankingan, seperti pada sistem pencarian informasi dan rekomendasi. MAP mengukur seberapa baik sistem menempatkan item relevan di posisi atas dalam daftar hasil. Untuk setiap query, dihitung Average Precision (AP), yaitu rata-rata precision pada setiap posisi di mana item relevan ditemukan dalam daftar hasil. MAP kemudian diperoleh dengan menghitung rata-rata AP dari seluruh query. Rumus MAP adalah:

$$MAP = \frac{1}{Q} \sum_{q=1}^Q AP(q) \quad (2.9)$$

di mana Q adalah jumlah query, dan $AP(q)$ adalah Average Precision untuk query ke- q .

MAP sangat berguna untuk mengevaluasi sistem di mana urutan hasil sangat penting, karena memberikan penalti jika item relevan muncul di posisi bawah. Nilai MAP berkisar antara 0 hingga 1, dengan nilai lebih tinggi menunjukkan performa ranking yang lebih baik. MAP banyak digunakan dalam penelitian internasional, misalnya pada bidang information retrieval dan sistem rekomendasi. [26]

U I M N
U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A