

**IMPLEMENTASI HYBRID CONTENT-BASED FILTERING PADA
SISTEM REKOMENDASI SKINCARE MENGGUNAKAN
RESNET-50 DAN TF-IDF**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

RIZALDI YANUAR

22.12.2453

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**IMPLEMENTASI HYBRID CONTENT-BASED FILTERING PADA
SISTEM REKOMENDASI SKINCARE MENGGUNAKAN
RESNET-50 DAN TF-IDF**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

RIZALDI YANUAR

22.12.2453

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2026**

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI HYBRID CONTENT-BASED FILTERING PADA
SISTEM REKOMENDASI SKINCARE MENGGUNAKAN
RESNET-50 DAN TF-IDF**

yang disusun dan diajukan oleh

Rizaldi Yanuar

22.12.2453

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 19 Mei 2026

Dosen Pembimbing,



M. Nuraminudin, S.Kom., M.Kom.

NIK. 190302408

HALAMAN PENGESAHAN
SKRIPSI
IMPLEMENTASI HYBRID CONTENT-BASED FILTERING PADA
SISTEM REKOMENDASI SKINCARE MENGGUNAKAN
RESNET-50 DAN TF-IDF

yang disusun dan diajukan oleh

Rizaldi Yanuar

22.12.2453

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Mei 2026

Susunan Dewan Penguji

Nama Penguji

Yoga Pristyanto, S.Kom., M.Eng.
NIK. 190302412

Atik Nurmasani, S.Kom., M.Kom.
NIK. 190302354

M. Nuraminudin, S.Kom., M.Kom.
NIK. 190302408

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 19 Mei 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, S.Kom., M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rizaldi Yanuar
NIM : 22.12.2453

Menyatakan bahwa Skripsi dengan judul berikut:

**IMPLEMENTASI HYBRID CONTENT-BASED FILTERING PADA
SISTEM REKOMENDASISKINCARE MENGGUNAKAN
RESNET-50 DAN TF-IDF**

Dosen Pembimbing : M. Nuraminudin, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 19 Mei 2026

Yang Menyatakan,



Rizaldi Yanuar

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa, yang telah melimpahkan rahmat, hidayah, serta karunia-Nya, sehingga penulis dapat menyelesaikan laporan tugas akhir ini dengan lancar. Tugas akhir ini disusun untuk memenuhi salah satu syarat kelulusan Program Sarjana pada Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Selain itu penulis dengan segala kerendahan hati ingin menyampaikan rasa terima kasih kepada semua pihak yang telah berjasa memberikan dukungan dan bantuan untuk menyelesaikan laporan tugas akhir ini. Untuk itu penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta
2. Ibu Prof. Dr. Kusrini, S.Kom., M.Kom., selaku Dekan Program Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta
3. Bapak Anggit Dwi Hartanto, M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas AMIKOM Yogyakarta
4. Bapak M. Nuraminudin, S.Kom., M.Kom., selaku dosen pembimbing yang memberikan arahan, saran, dan motivasi terhadap penulis
5. Kedua orang tua, keluarga besar, dan teman-teman tercinta yang memberikan semangat dan doa kepada penulis.

Yogyakarta, 19 Mei 2026

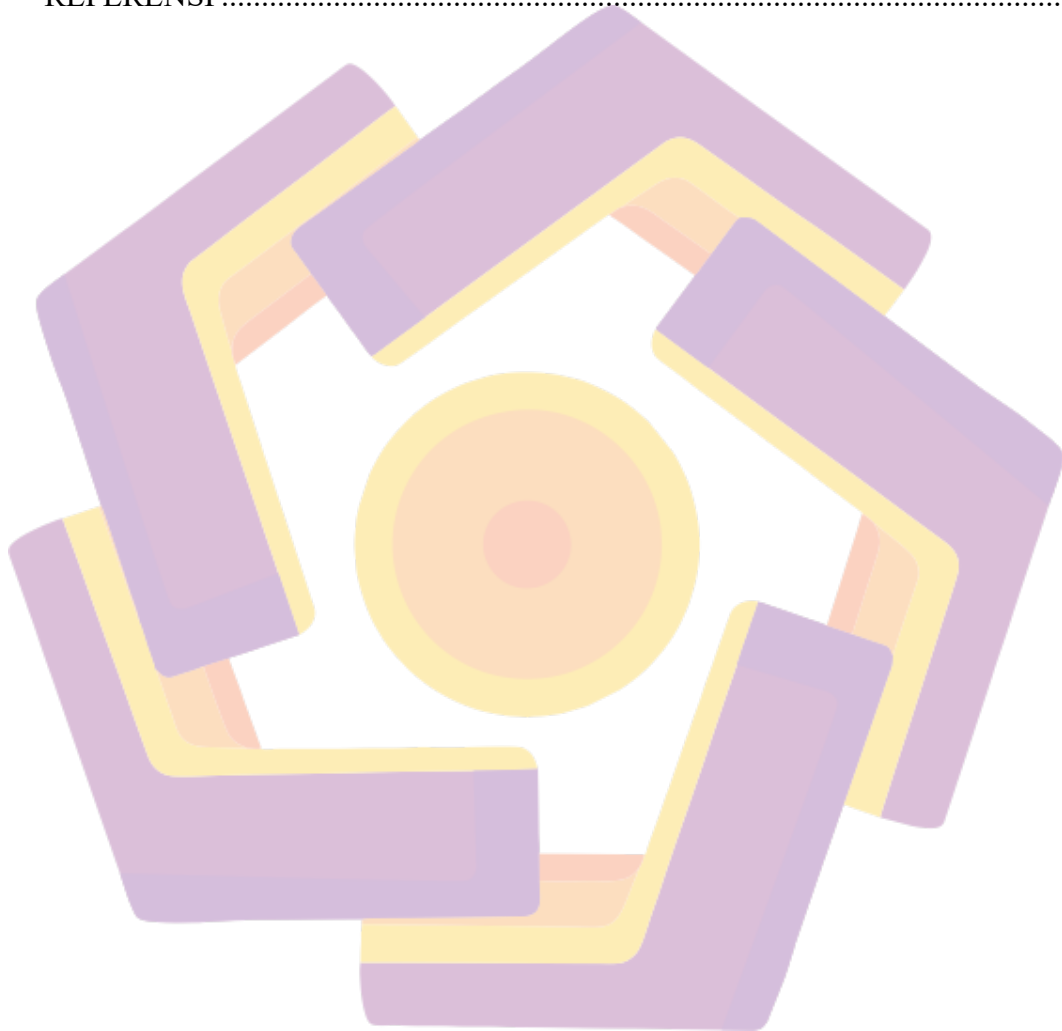
Rizaldi Yanuar

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMBANG DAN SINGKATAN	xi
DAFTAR ISTILAH	xii
INTISARI	xiv
ABSTRACT.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1 Studi Literatur	7
2.2 Dasar Teori.....	12
2.2.1 Sistem Rekomendasi	12
2.2.2 Hybrid Content-Based Filtering.....	13
2.2.3 Text Mining dan Preprocessing	14
2.2.4 TF-IDF	15
2.2.5 Convolutional Neural Network (CNN)	17
2.2.6 ResNet50.....	18
2.2.7 Cosine Similarity	19

2.2.8 Evaluasi Sistem	21
a. Precision@K	21
b. Silhouette Score	22
c. Visualisasi t-SNE	23
BAB III METODE PENELITIAN	25
3.1 Objek Penelitian	25
3.2 Alur Penelitian	32
3.3 Alat dan Bahan	35
3.3.1 Perangkat Keras (Hardware)	35
3.3.2 Perangkat Lunak (Software)	35
3.3.3 Bahan Penelitian	36
BAB IV HASIL DAN PEMBAHASAN	37
4.1 Pengumpulan Data	37
4.2 Preprocessing Data	38
4.2.1 Data Cleaning	38
a. Handling Duplicates	38
b. Handling Missing Values	39
4.2.2 Text Preprocessing	40
a. Case Folding	40
b. Cleansing	41
c. Whitespace Normalization	42
4.2.3 Image Preprocessing	43
4.3 Ekstraksi Fitur	44
4.3.1 ResNet 50	44
4.3.2 TF-IDF	47
4.4 Cosine Similarity	49
4.5 Sistem Rekomendasi	53
4.6 Evaluasi	55
4.6.1 Precision@5	56
4.6.2 Silhouette Score	57
4.6.3 Visualisasi t-SNE	58

4.6.4 Kesimpulan Evaluasi.....	60
4.7 Deployment.....	60
BAB V PENUTUP	65
5.1 Kesimpulan	65
5.2 Saran	66
REFERENSI	67

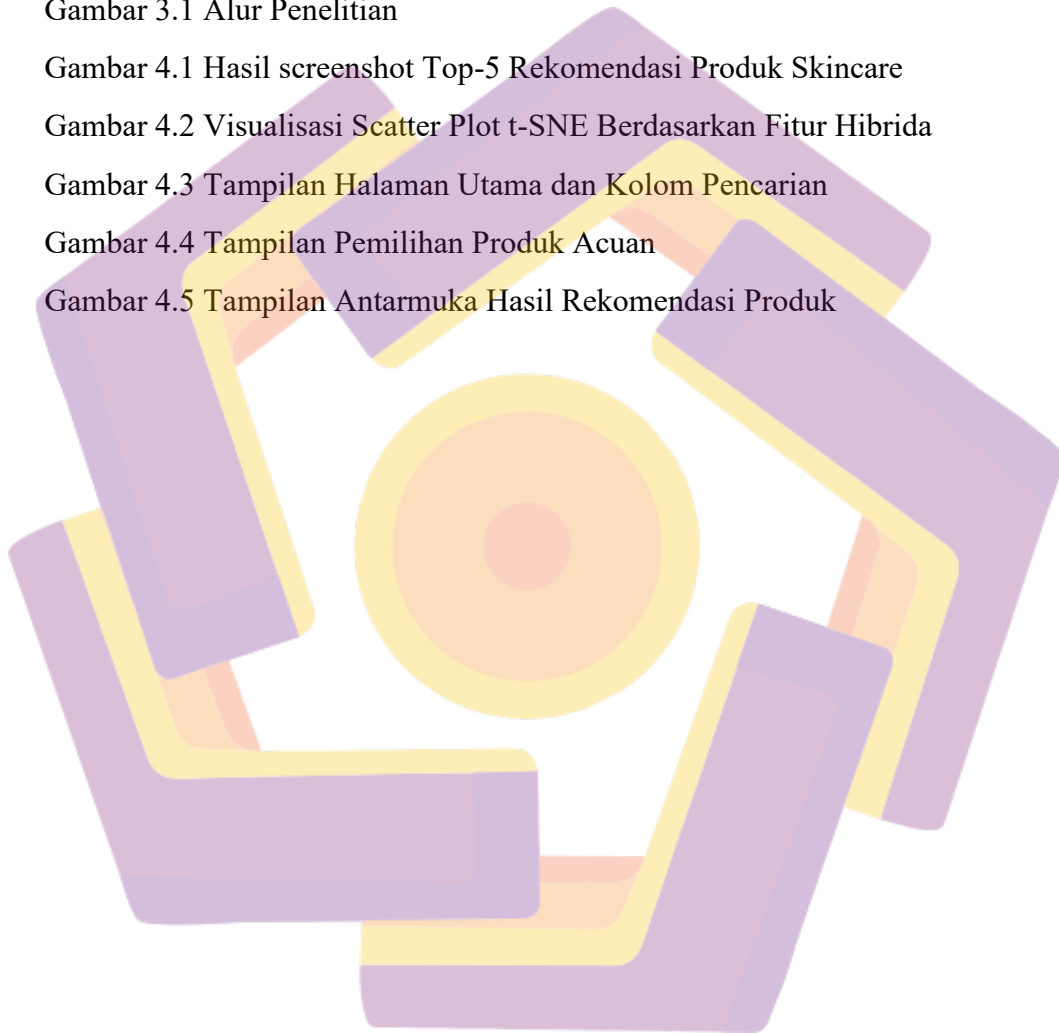


DAFTAR TABEL

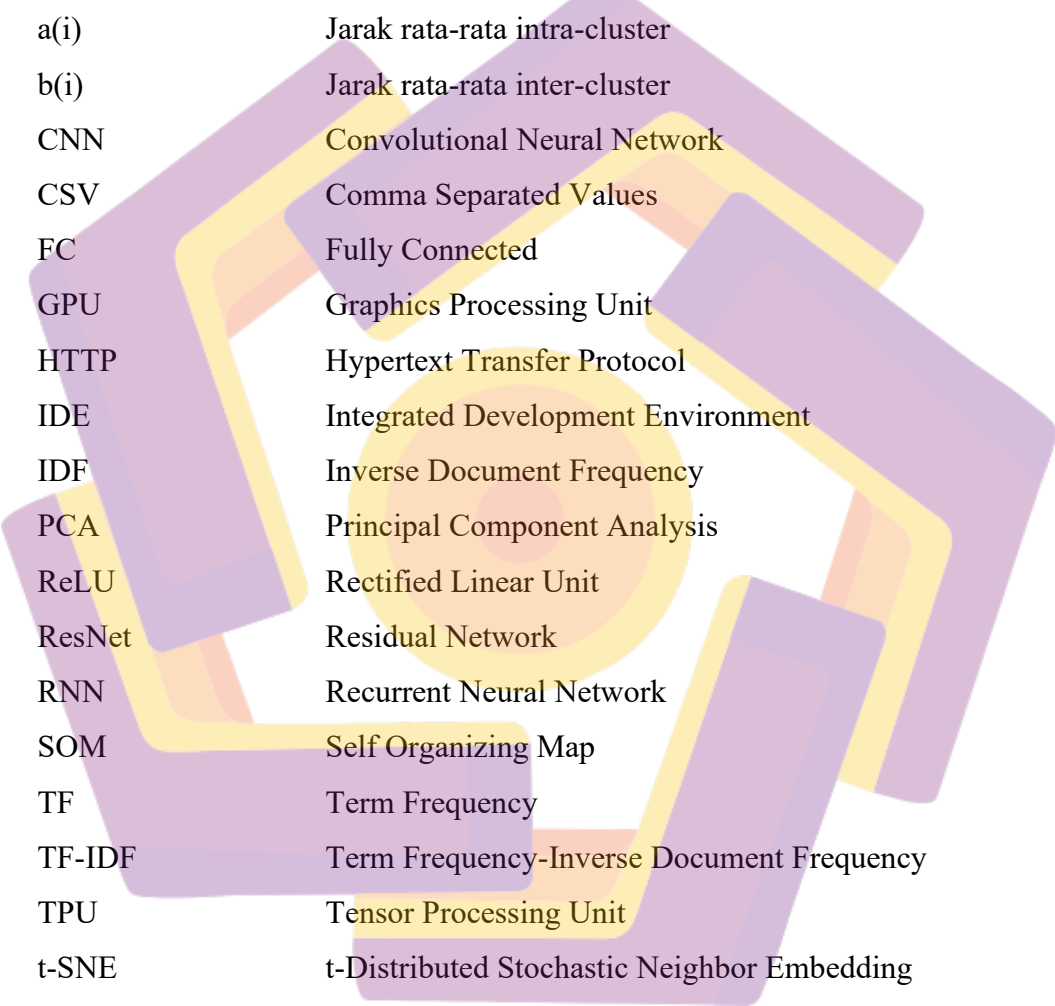
Tabel 2.1 Keaslian Penelitian	9
Tabel 3.1 Sampel Data Objek Penelitian	25
Tabel 4.1 Struktur Atribut Data	37
Tabel 4.2 Hasil Handling Duplicates	39
Tabel 4.3 Hasil Missing Values	39
Tabel 4.4 Hasil Implementasi Tahap Case Folding	40
Tabel 4.5 Hasil Implementasi Tahap Cleansing	41
Tabel 4.6 Hasil Whitespace Normalization	42
Tabel 4.7 Bukti Akuisisi dan Validasi Lokasi Citra Lokal	44
Tabel 4.8 Hasil Implementasi Ekstraksi Fitur Citra menggunakan ResNet50	45
Tabel 4.9 Hasil Ekstraksi Bobot TF-IDF pada 10 Kata Pertama Ingredients	48
Tabel 4.10 Hasil Komputasi Cosine Similarity dan Pembobotan Hybrid	51
Tabel 4.11 Hasil Top-5 Rekomendasi Produk Skincare (Kueri: "Serum")	54
Tabel 4.12 Hasil Pengujian Precision@5	56
Tabel 4.13 Hasil Pengujian Silhouette Score	58

DAFTAR GAMBAR

Gambar 2.1 Content-Based Filtering	13
Gambar 2.2 Convolutional Neural Network (CNN)	18
Gambar 2.3 Contoh Visualisasi t-SNE	24
Gambar 3.1 Alur Penelitian	33
Gambar 4.1 Hasil screenshot Top-5 Rekomendasi Produk Skincare	54
Gambar 4.2 Visualisasi Scatter Plot t-SNE Berdasarkan Fitur Hibrida	59
Gambar 4.3 Tampilan Halaman Utama dan Kolom Pencarian	61
Gambar 4.4 Tampilan Pemilihan Produk Acuan	62
Gambar 4.5 Tampilan Antarmuka Hasil Rekomendasi Produk	63



DAFTAR LAMBANG DAN SINGKATAN



θ	Sudut yang terbentuk antara dua vektor
$W_{t,d}$	Bobot akhir kata pada dokumen
$\text{Sim}(A, B)$	Skor kemiripan (Cosine Similarity)
S_i	Nilai Silhouette Score
$a(i)$	Jarak rata-rata intra-cluster
$b(i)$	Jarak rata-rata inter-cluster
CNN	Convolutional Neural Network
CSV	Comma Separated Values
FC	Fully Connected
GPU	Graphics Processing Unit
HTTP	Hypertext Transfer Protocol
IDE	Integrated Development Environment
IDF	Inverse Document Frequency
PCA	Principal Component Analysis
ReLU	Rectified Linear Unit
ResNet	Residual Network
RNN	Recurrent Neural Network
SOM	Self Organizing Map
TF	Term Frequency
TF-IDF	Term Frequency-Inverse Document Frequency
TPU	Tensor Processing Unit
t-SNE	t-Distributed Stochastic Neighbor Embedding
URL	Uniform Resource Locator

DAFTAR ISTILAH



Case Folding	Langkah awal penyeragaman teks dengan mengubah semua huruf kapital menjadi huruf kecil.
Cleansing	Proses membuang karakter-karakter derau (seperti angka dan simbol) yang tidak informatif dari data teks.
Cold-start	Kendala pada sistem rekomendasi ketika produk baru belum memiliki riwayat interaksi atau penilaian pengguna.
Content-Based Filtering	Metode pemberian rekomendasi berdasarkan kemiripan atribut spesifik dari sebuah item.
Cosine Similarity	Metode untuk menghitung tingkat kemiripan antara dua vektor dengan mengukur nilai kosinus sudut di antara keduanya.
Deep Learning	Cabang algoritma pembelajaran mesin berbasis jaringan saraf tiruan, seperti arsitektur ResNet50.
Deployment	Tahapan akhir implementasi sistem ke dalam bentuk aplikasi berbasis web agar dapat digunakan secara interaktif.
Hybrid Multimodal	Pendekatan sistem yang mengintegrasikan berbagai jenis representasi fitur sekaligus, yaitu visual (citra) dan tekstual (teks).
Ingredients	Daftar teks komposisi zat kimia dan bahan alami yang menjadi penyusun suatu produk perawatan kulit.
Information Overload	Fenomena kelebihan informasi yang membingungkan konsumen akibat terlalu banyaknya variasi produk di pasaran.

Precision@K

Metrik evaluasi untuk mengukur persentase ketepatan atau relevansi kategori dari rekomendasi pada peringkat teratas.

Preprocessing

Tahapan pengolahan data mentah untuk meminimalkan derau (noise) dan menyeragamkan format sebelum diolah algoritma.

ResNet50

Arsitektur Convolutional Neural Network dengan kedalaman 50 lapisan untuk mengekstraksi fitur visual gambar.

Silhouette Score

Metrik evaluasi yang mengukur validitas dan kualitas batas pengelompokan (klaster) data di dalam ruang vektor.

Term Frequency (TF)

Metrik yang mengukur seberapa sering sebuah kata (term) muncul dalam satu dokumen tertentu.

TF-IDF

Metode pembobotan statistik untuk merepresentasikan tingkat kepentingan sebuah kata dalam sebuah dokumen teks.

t-SNE

Algoritma reduksi dimensi untuk memvisualisasikan persebaran data berdimensi tinggi ke dalam grafik dua dimensi.

INTISARI

Perkembangan pesat industri perawatan kulit telah menyebabkan membanjirnya variasi produk di pasaran yang berdampak pada kebingungan konsumen dalam memilih produk yang sesuai preferensinya. Fenomena kelebihan informasi ini menyulitkan pengguna untuk menemukan produk alternatif yang memiliki kemiripan karakteristik, baik dari segi kandungan maupun kemasan, terutama pada produk baru yang belum memiliki riwayat penilaian dari pengguna lain. Penelitian ini bertujuan menyelesaikan masalah tersebut dengan membangun sistem rekomendasi menggunakan pendekatan Hybrid Multimodal Content Based Filtering. Metode ini mengintegrasikan ekstraksi fitur visual dari citra kemasan produk menggunakan algoritma Convolutional Neural Network dengan arsitektur ResNet dan ekstraksi fitur tekstual dari komposisi bahan menggunakan metode pembobotan Term Frequency Inverse Document Frequency. Kedua fitur tersebut digabungkan untuk menghitung tingkat kemiripan antar produk menggunakan algoritma Cosine Similarity. Hasil akhir dari penelitian ini adalah sebuah sistem yang mampu memberikan rekomendasi produk perawatan kulit yang relevan berdasarkan kemiripan visual dan kandungan bahan secara bersamaan. Evaluasi kinerja sistem dilakukan menggunakan metrik presisi kategori dan validasi kluster untuk memastikan kualitas pengelompokan data tanpa label. Penelitian ini memberikan kontribusi nyata bagi konsumen dalam mempermudah pencarian produk yang akurat serta membantu mengatasi kendala rekomendasi pada produk baru yang belum memiliki interaksi pengguna sebelumnya.

Kata kunci: Sistem Rekomendasi, Skincare, Hybrid Multimodal, TF-IDF, ResNet50.

ABSTRACT

The rapid growth of the skincare industry has led to an influx of product variations in the market, resulting in consumer confusion when selecting products that suit their preferences. This phenomenon of information overload complicates the process for users to discover alternative products sharing similar characteristics, both in terms of ingredients and packaging, particularly for new products lacking a history of user ratings. This study aims to address this issue by developing a recommendation system utilizing a Hybrid Multimodal Content-Based Filtering approach. This method integrates visual feature extraction from product packaging images using the Convolutional Neural Network algorithm with ResNet architecture and textual feature extraction from ingredient compositions using the Term Frequency-Inverse Document Frequency weighting method. Both features are combined to calculate the similarity level between products using the Cosine Similarity algorithm. The final outcome of this research is a system capable of providing relevant skincare product recommendations based on visual similarity and ingredient content simultaneously. System performance evaluation is conducted using category precision metrics and cluster validation to ensure the quality of unlabeled data grouping. This study contributes significantly to consumers by facilitating accurate product searches and assisting in overcoming recommendation challenges for new products that lack prior user interaction.

Keyword: Recommendation System, Skincare, Hybrid Multimodal, TF-IDF, ResNet50.