

Sistem Rekomendasi Produk Somethinc Menggunakan Metode *Content-based Filtering*

Nailatul Azizah^a, Anief Fauzan Rozi^b

^aSistem Informasi, Fakultas Teknologi Informasi, Universitas Mercu Buana Yogyakarta, 201210010@student.mercubuana-yogya.ac.id

^bSistem Informasi, Fakultas Teknologi Informasi, Universitas Mercu Buana Yogyakarta, anief@mercubuana-yogya.ac.id

Submitted: 26-05-2024, Reviewed: 06-06-2024, Accepted 24-06-2024

<https://doi.org/10.47233/jteksis.v6i3.1411>

Abstract

Many consumers face difficulties in choosing the right skincare products due to a lack of understanding of skin types and available product information. Therefore, there is a need for a system that can recommend skincare products tailored to individual consumer needs. This study aims to develop an application that provides skincare product recommendations based on previously used products. The method applied is content-based filtering, which compares the content of various products to generate recommendation rankings. The minimum support and confidence values are calculated to determine association rules for itemset combinations. The cosine similarity algorithm is used to measure the similarity between words, while product descriptions are analyzed using TF-IDF (Term Frequency-Inverse Document Frequency). The similarity weights are calculated using the cosine similarity algorithm. The results of the study show that the cosine similarity values calculated using 8 skincare product data points indicate the highest similarity value is 0.722 for two product documents and the lowest value is 0 for one product document. This application can provide relevant skincare product recommendations that match user preferences.

Keywords: Recommendation Systems, Content-Based Filtering, Cosine Similarity, TF-IDF Weighting.

Abstrak

Banyak konsumen menghadapi kesulitan dalam pemilihan produk perawatan kulit yang sesuai karena kurangnya pemahaman tentang jenis kulit dan informasi produk yang tersedia, oleh karena itu, diperlukan sistem yang mampu merekomendasikan produk skincare yang sesuai dengan kebutuhan individu konsumen. Penelitian ini bertujuan mengembangkan aplikasi yang memberikan rekomendasi produk skincare berdasarkan produk yang telah digunakan sebelumnya. Metode yang diterapkan adalah *content-based filtering*, yaitu dengan membandingkan konten berbagai produk untuk menghasilkan peringkat rekomendasi. Nilai minimum *support* dan *confidence* dihitung untuk menentukan aturan asosiasi kombinasi itemset. Algoritma *cosine similarity* digunakan untuk mengukur kemiripan antar kata, sementara deskripsi produk dianalisis menggunakan TF-IDF (Term Frequency-Inverse Document Frequency). Bobot kemiripan dihitung dengan algoritma *cosine similarity*. Hasil penelitian menunjukkan nilai *cosine similarity* dari perhitungan dengan menggunakan 8 data produk skincare dengan nilai kemiripan tertinggi yaitu 0,722 pada 2 dokumen produk dan nilai terendah 0 pada satu dokumen produk, aplikasi ini dapat memberikan rekomendasi produk skincare yang relevan dan sesuai dengan preferensi pengguna.

Keywords: Sistem Rekomendasi, Content-Based Filtering, Cosine Similarity, Pembobotan TF-IDF.

This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license



PENDAHULUAN

Kemajuan teknologi saat ini memperlihatkan peningkatan karena teknologi berkembang bersamaan dengan perkembangan ilmu pengetahuan [1]. Perkembangan teknologi saat ini berdampak signifikan pada kemajuan industri kecantikan yang juga berkembang pesat [2]. Menurut informasi yang diperoleh dari Kementerian Perindustrian, jumlah industri kosmetik dan usaha kecil menengah (IKM) yang signifikan di Indonesia pada tahun 2019 mencapai 797 perusahaan, meningkat dari 760 perusahaan pada tahun sebelumnya. Data dari BKPM (Badan Koordinasi Penanaman Modal) menyatakan bahwa dari total 797 industri kosmetik nasional, sebanyak 294 di antaranya terdaftar pada tahun yang sama. Prediksi menunjukkan bahwa permintaan akan produk perawatan kulit terus

menunjukkan peningkatan dari tahun ke tahun. Tidak hanya itu, menurut Statista, terdapat peningkatan sebesar 7,2% dalam penggunaan skincare dan kosmetik pada tahun 2022 dibandingkan dengan tahun-tahun sebelumnya [3].

Kosmetik dibagi menjadi dua kategori. Pertama, kosmetik perawatan kulit (*skin care*), yang berfungsi untuk merawat, memelihara, dan menjaga kesehatan kulit. Kedua, kosmetik riasan (*make up*), yang berguna untuk mempercantik dan menyamarkan ketidaksempurnaan pada wajah, sehingga meningkatkan penampilan [4]. Setiap orang, terutama wanita, bermimpi memiliki kulit yang sehat dan cantik. Salah satu cara untuk merawat kebersihan dan kesehatan kulit adalah melalui penggunaan skincare atau produk perawatan kulit. Melalui perawatan ini, kebersihan dan

kesehatan kulit dapat dipertahankan, sehingga kulit terlihat sehat, segar, dan menarik bagi yang melihatnya. *Skincare* adalah kegiatan merawat kulit dengan menggunakan bahan-bahan khusus, terutama untuk wajah. Penting untuk mengetahui kondisi dan masalah kulit wajah sebelum memilih produk *skincare* yang akan digunakan.

SOMETHINC menawarkan beragam jenis produk *make up* dan perawatan kulit, namun hal ini sering kali membuat para pelanggan bingung dalam memilih produk yang sesuai dengan kebutuhan mereka [5]. Sebelumnya, pelanggan biasanya melakukan pencarian produk secara manual melalui internet. Agar pelanggan dapat menemukan lebih banyak pilihan produk yang sesuai dengan minat dan kebutuhan mereka, diperlukan sistem rekomendasi yang mampu memberikan saran berdasarkan preferensi dan kebutuhan pelanggan.

Dalam penyusunan penelitian ini, penting untuk merujuk pada penelitian sebelumnya untuk memperluas pengetahuan dan wawasan peneliti serta menghindari duplikasi penelitian. Beberapa penelitian terdahulu menjadi referensi utama.

Penelitian Safitri, Helilintar, & Wahyuniar (2021) bertujuan mengembangkan aplikasi rekomendasi produk *skincare* berdasarkan produk yang disukai sebelumnya. Metode yang digunakan pada penelitian ini adalah *Content-Based Filtering* dan Algoritma Apriori, yang dihitung dari nilai kemiripan konten untuk rating produk dan mengevaluasi minimum support dan *confidence* untuk aturan asosiasi. Hasilnya menunjukkan rating tertinggi 0,447 dan nilai *confidence* 88,89% [6].

Penelitian kedua Alkaff, Khatimi, and Eriady (2020) melakukan penelitian kedua yang membahas penerapan sistem rekomendasi menggunakan metode Filter Berbasis Konten untuk memberikan rekomendasi buku berdasarkan kesamaan fitur-fitur pada buku tersebut dengan menggunakan *Weighted Tree Similarity*. Tujuan dari penelitian ini adalah untuk memudahkan pengunjung Perpustakaan Daerah Kalimantan Selatan dalam menemukan buku sesuai dengan preferensi mereka. Hasil penelitian menunjukkan bahwa presisi dari lima skenario pengujian mencapai 88% [7].

Penelitian ketiga oleh Wijaya & Alfian (2018) menggunakan dua pendekatan, yaitu *content-based filtering* dan *collaborative filtering*, untuk memberikan rekomendasi laptop kepada pelanggan berdasarkan minat dan kebutuhan mereka. Dalam metode *collaborative filtering*, mereka menerapkan algoritma *Adjusted-cosine similarity* untuk menghitung kemiripan antar pelanggan, serta menggunakan algoritma *weighted sum* untuk perhitungan prediksi. Sementara itu, dalam metode *content-based filtering*, mereka menggunakan algoritma *tf-idf* untuk mengevaluasi ketersediaan konten. [8].

Fiarni, Maharani, dan Calista (2019) melakukan penelitian keempat yang mengeksplorasi sistem rekomendasi untuk perusahaan pewarna tekstil. Penelitian ini bertujuan untuk merekomendasikan produk baru atau alternatif kepada calon pelanggan berdasarkan karakteristik individu dan riwayat transaksi mereka. Metode yang digunakan melibatkan penggunaan *cosine similarity* untuk menilai kesamaan produk dan *content-based filtering* untuk menilai model rekomendasi pelanggan. Hasil penelitian menunjukkan tingkat kesamaan sebesar 93,47% [9].

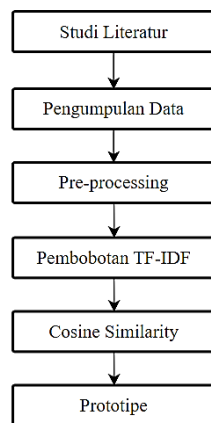
Penelitian ini menggunakan metode penyaringan berbasis konten (*Content-based filtering*) yang memanfaatkan informasi produk untuk memberikan rekomendasi. Metode ini dipilih karena tidak memerlukan penilaian dari pengguna lain sebagai dasar rekomendasi. Dalam penelitian ini, deskripsi produk yang sesuai dengan kata kunci yang dicari oleh pelanggan dijadikan sebagai parameter. Berdasarkan produk yang dicari oleh pelanggan, sistem dapat merekomendasikan produk baru yang memiliki deskripsi konten yang mirip dengan produk yang telah dilihat sebelumnya. Inovasi dari penelitian ini adalah penerapan algoritma pemrosesan bahasa alami (*Natural Language Processing*) terbaru untuk menganalisis dan memahami deskripsi produk secara lebih baik. Ini memungkinkan sistem memberikan rekomendasi yang lebih akurat dan relevan berdasarkan pemahaman yang lebih mendalam terhadap konten produk.

METODE PENELITIAN

Data dalam penelitian ini adalah data sekunder yang berasal dari sumber tidak langsung, yang diperoleh melalui perantara seperti orang lain atau dokumen dan studi literatur pada jurnal atau pun buku yang membahas tentang *skincare* dan *make up*.

2.1. Jalan Penelitian

Penelitian ini melibatkan enam langkah utama, yang mencakup studi literatur, pengumpulan data, *pre-processing*, pembobotan TF-IDF, *cosine similarity*, dan *prototipe* seperti pada Gambar 1.



Gambar 1. Alur Penelitian

- a. **Studi Literatur.**
Pada studi literatur ini, dilakukan pengumpulan, pencarian, dan memahami informasi, serta informasi yang sesuai dengan penelitian ini.
- b. **Pengumpulan Data.**
Dalam penelitian ini, pengumpulan data dilakukan melalui studi pustaka, di mana data yang diperlukan berasal dari referensi-referensi yang diperoleh dari penelitian-penelitian sebelumnya [10]. Proses pengumpulan data ini berkaitan dengan akuisisi pengetahuan tentang *skincare* dan *make-up*. Langkah akuisisi pengetahuan ini melibatkan pengumpulan informasi tentang masalah-masalah yang akan diteliti, yang kemudian akan menjadi panduan dalam pengembangan sistem rekomendasi.
- c. **Pre-processing.**
Pre-processing teks adalah proses yang melibatkan beberapa teknik, memecah teks menjadi unit-unit seperti kata-kata, menghilangkan kata-kata yang tidak penting, dan mengubah kata-kata menjadi bentuk dasarnya, semua ini dilakukan untuk membuat teks menjadi lebih terstruktur dan siap untuk analisis lebih lanjut.
- d. **Pembobotan TF-IDF.**
Metode TF-IDF adalah suatu teknik yang digunakan untuk menentukan seberapa pentingnya sebuah kata (*term*) dalam konteks sebuah dokumen. Teknik ini mengintegrasikan dua konsep, yaitu frekuensi kemunculan kata dalam dokumen tertentu dan kebalikan dari frekuensi dokumen yang memiliki kata tersebut [11].
- e. **Cosine Similarity.**
Cosine similarity adalah teknik yang dipakai untuk mengukur tingkat kemiripan antar dua objek. Kesamaan dihitung antara

kueri dan dokumen. Nilai kemiripan didapat dengan menghitung nilai kemiripan antara dua vektor. Ukuran kemiripan digunakan untuk menentukan tingkat kemiripan antara data [12].

f. **Prototipe.**

Prototipe dibuat sebagai sebuah model yang bisa diakses oleh pengguna ketika mereka memasukkan kata kunci pada aplikasi sistem rekomendasi, dengan maksud untuk meningkatkan kenyamanan pengguna dan meningkatkan relevansi rekomendasi yang diberikan.

2.2 Sistem Rekomendasi

Sistem rekomendasi adalah sistem yang menyajikan informasi dan rekomendasi kepada pengguna untuk membantu mereka membuat keputusan, dengan menggunakan data historis [13]. Sistem rekomendasi menyajikan daftar item yang sesuai dengan individu pengguna. Sistem ini menganalisis data pengguna, informasi pribadi, dan konteks pengguna untuk menentukan preferensi mereka. Dengan sistem rekomendasi menyaring informasi secara personal sehingga hanya objek yang sesuai dengan pengguna yang ditampilkan menggunakan model rekomendasi [14].

2.3 Skincare

Istilah "*skincare*" berasal dari bahasa Inggris yang mengacu pada perawatan kulit. Secara umum, *skincare* merujuk pada rangkaian perawatan kulit yang melibatkan penggunaan berbagai produk untuk menjaga kesehatan dan kecantikan kulit. Nutrisi yang penting bagi kulit dapat dipenuhi baik dari dalam tubuh melalui makanan dan minuman yang dikonsumsi, maupun dari luar tubuh melalui produk *skincare*. Langkah yang krusial dalam menjaga kesehatan dan kecantikan kulit wajah adalah memastikan kulit bebas dari sel-sel kulit mati, debu, kotoran, dan sisa-sisa makeup yang menempel, sehingga juga membantu mencegah berbagai masalah kulit [15].

2.4 Pengertian Metode *Content-based filtering*

Content-based filtering tidak mempertimbangkan kontribusi dari pengguna lain dalam memberikan rekomendasi, melainkan hanya bergantung pada preferensi pengguna tersebut [16]. Algoritma ini memilih item dengan konten yang mirip berdasarkan apa yang dicari oleh pengguna, tanpa memperhatikan penilaian dari pengguna lain [17]. Meskipun hal ini dapat menghasilkan rekomendasi yang kurang bervariasi, *Content-based filtering* tetap berfungsi dengan baik bahkan jika pengguna tidak memberikan penilaian [18].

Metode *Content-based filtering* mengekstrak informasi dari item yang akan direkomendasikan,

lalu membandingkannya dengan informasi dari produk yang telah dikunjungi. Metode ini sering digunakan untuk merekomendasikan artikel, atau situs web. Beberapa metode yang sering digunakan dalam penyaringan berbasis konten meliputi penggunaan TF-IDF, Klasifikasi Bayes, analisis klaster, pohon keputusan, dan jaringan saraf tiruan. Kelebihan dari metode *Content-based filtering* adalah kemampuannya memberikan rekomendasi tanpa memerlukan penilaian dari pengguna, hanya dengan mempertimbangkan informasi dari item itu sendiri.

2.5 TF-IDF (Term Frequency Inverse Document Frequency)

Secara garis besar, TF-IDF adalah sebuah metode yang mengukur relevansi sebuah *term* dalam sebuah dokumen dengan membandingkan frekuensi kemunculannya dalam dokumen tersebut dengan seberapa sering kemunculan kata tersebut dalam seluruh dokumen [19]. Metode TF-IDF adalah salah satu metode pembobotan term yang banyak digunakan [20]. Kata-kata dengan nilai TF yang tinggi menunjukkan pentingnya kata tersebut dalam dokumen, sementara DF mengindikasikan seberapa umum kata tersebut muncul dalam seluruh koleksi dokumen. Metode ini menggunakan perhitungan kemunculan kata dalam banyak dokumen [21].

Dalam penelitian ini, digunakan algoritma TF-IDF dengan dua jenis perhitungan yaitu TF biner dan TF murni [22]. TF biner berfungsi untuk menentukan keberadaan atau tidaknya sebuah kata dalam sebuah dokumen, di mana nilai satu (1) menunjukkan keberadaan dan nol (0) menunjukkan ketiadaan. Sedangkan TF murni menggambarkan jumlah kemunculan kata dalam sebuah dokumen. Sebagai contoh, jika sebuah kata muncul lima kali dalam dokumen, maka nilainya adalah lima (5).

IDF merupakan langkah perhitungan setelah nilai TF diperoleh. IDF mengukur seberapa luas distribusi TF pada seluruh koleksi dokumen yang relevan dengan produk yang bersangkutan. Berikut adalah rumus 1 dan 2 yang digunakan pada penelitian ini:

- 1) $IDF = \left(\frac{D}{DF} \right)$
- 2) $W = TF * (IDF + 1)$

TF : Frekuensi sebuah kata atau istilah muncul dalam sebuah dokumen.

IDF : Kebalikan dari frekuensi dokumen, yang dihitung dengan rumus logaritma dari jumlah total dokumen dibagi dengan jumlah dokumen yang mengandung kata tersebut (*term*).

D : Total jumlah dokumen.

DF: Jumlah dokumen yang mengandung kata atau istilah tersebut.

W: Bobot diberikan untuk setiap dokumen.

Rumus (1) adalah cara untuk menghitung *Inverse Document Frequency* (IDF), yang menunjukkan pentingnya sebuah kata dengan mempertimbangkan seberapa sering kata tersebut muncul dalam dokumen dibandingkan dengan total jumlah dokumen dalam koleksi tersebut [23]. Rumus (2), di sisi lain, adalah cara untuk menghitung bobot suatu kata dalam model TF-IDF setelah mempertimbangkan frekuensi kemunculan kata tersebut dalam dokumen. Dalam rumus tersebut, frekuensi kemunculan kata dikalikan dengan nilai IDF yang telah dimodifikasi dengan penambahan 1, kemungkinan untuk menghindari pembagian oleh nol atau memberikan bobot tambahan kepada kata-kata yang muncul di semua dokumen[24]. Ini membantu menentukan seberapa pentingnya sebuah kata dalam sebuah dokumen dalam konteks seluruh koleksi dokumen[25].

HASIL DAN PEMBAHASAN

Proses penentuan rekomendasi produk something dilakukan melalui serangkaian langkah. Pertama, teks yang terkait dengan produk dipersiapkan melalui proses pra-pemrosesan, yang meliputi langkah-langkah seperti pembersihan teks dan penghapusan kata-kata yang tidak relevan. Selanjutnya, bobot diberikan pada setiap kata dalam teks menggunakan metode TF-IDF, yang membantu dalam menilai kepentingan relatif dari kata-kata tersebut dalam konteks keseluruhan. Kemudian, untuk menentukan kesamaan antara kata kunci yang dimasukkan oleh pengguna dan teks produk, rumus kesamaan kosinus digunakan [26]. Proses ini mengukur sudut antara vektor kata kunci pengguna dan vektor representasi teks produk dalam ruang fitur. Jika sudutnya mendekati nol, maka kedua vektor tersebut dianggap mirip, dan produk tersebut dianggap sebagai rekomendasi yang potensial. Dengan demikian, proses rekomendasi produk ini memungkinkan pengguna untuk menemukan produk yang sesuai dengan kebutuhan mereka ketika memasukkan kata kunci seperti "serum untuk kulit berminyak".

3.1 Teks *Preprocessing*

Pemrosesan teks adalah proses penyaringan data teks agar menjadi lebih terstruktur melalui beberapa tahap. Hasil dari pemrosesan ini berupa daftar kata-kata penting yang memiliki arti signifikan dan dapat memberikan makna pada konten dokumen atau kumpulan dokumen [27].

Lakukan prapemrosesan teks pada deskripsi produk dengan maksud agar informasi yang digunakan bisa diubah menjadi nilai menggunakan TF-IDF dan kemiripan kosinus. Langkah-langkah dalam proses ini termasuk tokenisasi dan penghapusan kata penghubung. Prapemrosesan hanya dilakukan pada produk yang mengandung

kata kunci, sehingga pencarian menjadi lebih efisien dan cepat. Dengan demikian, deskripsi produk yang mencakup kata kunci dalam penelitian ini akan mengalami tahapan prapemrosesan.

Tabel 1. Data Preprocessing Produk

Kode	Nama Produk	Deskripsi
D1	Calm Down Skinpair R-Cover Cream	Pelembab vegan bertekstur ringan yang melindungi barier kulit cocok sebagai moisturizer untuk kulit sensitif dan memperkuat barier kulit meningkatkan kelembaban.
D2	SKIN GOALS Moisture Silk Creme	Pelembab gel vegan yang sangat ringan dengan kekuatan pencerah 5X, yang berfungsi sebagai pelembab untuk meningkatkan kelembaban kulit dan mengurangi flek hitam dalam periode 28 hari Produk pelembab ini dirancang khusus untuk kulit yang normal.
D3	CERAMIC SKIN Saviour Moisturizer Gel	Pelembab Gel Krim ini sangat lengkap dengan 10 Bahan Berkualitas Tinggi yang telah teruji untuk menjaga kelembaban kulit selama 24 jam Somethinc Ceramic Skin Saviour Moisturizer Gel cocok untuk merawat kulit yang beragam termasuk kulit kombinasi dan juga dapat berperan sebagai pelembab wajah untuk kulit yang mengalami kekeringan dan kekusaman.
D4	BEE PROPOLIS Honey Mask	Somethinc Sleeping Mask adalah produk perawatan kulit terbaik dengan formula Superfood Bee Glow yang membuat kulit terlihat bercahaya dan bebas kilap keesokan paginya. Cocok digunakan sebagai sleeping mask untuk jenis kulit yang berminyak.
D5	SKIN GOALS Moisture Silk Creme	Gel-cream vegan yang sangat ringan ini memiliki kekuatan pencerah 5X dan berfungsi sebagai pelembab untuk memperbaiki kulit dan mengurangi flek hitam dalam periode 28 hari Cocok digunakan untuk semua jenis kulit termasuk kulit sensitif dan berminyak.
D6	ACNEDOT Treatment Moisturizer Gel	Pelembab wajah vegan dengan kekuatan melawan jerawat 5X membantu mengatasi bakteri penyebab jerawat, mengurangi tampilan pori-pori, mengontrol produksi minyak berlebih, dan mengurangi jerawat dalam periode 28 hari.

D7	PEPTINOL Granactive Retinoid Peptide Night Moisturizer Creme	Memiliki BOTOX LIKE EFFECT yang membuat kulit terlihat lebih muda & segar. Cocok digunakan jenis kulit normal, sensitif dan berminyak.
D8	SOS Bakuchiol Electrolyte Rich Moisturizer Silk Creme	Memiliki teknologi Digital Stress Protector yang dapat melindungi kulit dari penuaan dini akibat paparan sinar Blue Light setiap hari. Cocok digunakan jenis normal.

3.2 Pembobotan TF-IDF

Deskripsi produk yang mengandung kata kunci diberi bobot 1 dalam proses pembobotan dokumen tersebut.

Tabel 2. Pembobotan Term

Term	TF	Pelembab	Sensitif	Minyak	Jerawat	kulit
Q	1	1	1	1	1	1
D1	1	1	0	0	0	1
D2	1	0	1	0	0	1
D3	1	0	0	0	0	1
D4	1	0	1	0	0	0
D5	0	1	1	1	1	1
D6	1	1	1	0	0	1
D7	0	1	1	1	1	0
D8	0	0	0	0	0	1

DF adalah jumlah dokumen yang mengandung term tertentu. Untuk menghitung nilai IDF, kita menggunakan rumus $IDF = \log(n/DF)$, di mana n adalah jumlah total dokumen dalam koleksi. Rumus ini memberikan bobot yang lebih tinggi pada term yang jarang muncul di dalam dokumen.

Tabel 3. Hasil IDF

Term	DF	D/DF	IDF
Pelembab	5	1.6	0.29
Sensitif	5	1.6	0.29
Berminyak	5	1.6	0.29
Jerawat	4	2	0.301
Kulit	7	1.14	0.08

Lakukan perhitungan untuk wdt menggunakan rumus $wdt = tf.idf$. Dengan cara mendapatkan bobot yang tepat untuk setiap term dalam dokumen berdasarkan seberapa sering mereka muncul (tf) dan seberapa jarang mereka muncul di seluruh koleksi dokumen (idf).

Tabel 4. Hasil Perhitungan Wdt

Term	Wdt=TF.IDF				
	Pelembab	Sensitif	Minyak	Jerawat	kulit
Q	0.29	0.29	0.29	0.301	0.08
D1	0.29	0.29	0	0	0.08
D2	0.29	0	0.29	0	0.08
D3	0.29	0	0	0	0.08
D4	0.29	0	0.29	0	0
D5	0	0.29	0.29	0.301	0.08
D6	0.29	0.29	0.29	0	0.08
D7	0	0.29	0.29	0.301	0
D8	0	0	0	0	0.08

Setiap dokumen dijadikan vektor dalam ruang fitur, dimana panjang vektor sesuai dengan jumlah kata-kata unik dalam kumpulan dokumen. Tiap komponen vektor merepresentasikan bobot dari kata-kata dalam dokumen tersebut, yang dihitung dengan nilai TF-IDF.

Tabel 5. Hasil Perhitungan Saklar

Term	WD*Wdi					
D1	0.841	0.841	0	0	0	0.0232
D2	0.0841	0	0.0841	0	0	0.0232
D3	0.841	0	0	0	0	0.0232
D4	0.0841	0	0.0841	0	0	0
D5	0	0.08729	0.08729	0.08729	0.0906	0.0241
D6	0.0232	0.0232	0.0232	0.0232	0	0.0064
D7	0	0.08729	0.08729	0.08729	0.0906	0
D8	0	0	0	0	0	0.064
D8	0	0	0	0	0	0.064

Panjang vektor untuk setiap dokumen dihitung dengan menggunakan formula *Euclidean norm*, yaitu akar kuadrat dari total kuadrat bobot kata-kata dalam dokumen. Panjang vektor ini memberikan indikasi tentang kompleksitas atau relevansi dokumen tersebut dalam kaitannya dengan kumpulan dokumen secara keseluruhan.

Tabel 6. Hasil Perhitungan Panjang Vektor

Term	Panjang Vektor
D1	0.119
D2	0.148
D3	0.148
D4	0.128
D5	0.067
D6	0.148
D7	0.128
D8	0.179

3.3 Cosine similarity

Cosine similarity adalah teknik yang digunakan untuk mengevaluasi seberapa mirip dua dokumen. Untuk membuat sistem peringkat

rekomendasi, digunakan *cosine similarity* antara dokumen pertanyaan (Q) dan setiap dokumen lainnya (D1-D8). Dokumen yang memiliki similaritas kosinus yang tinggi dengan dokumen pertanyaan akan mendapatkan peringkat yang lebih tinggi.

Cosine Similarity (Q, D1) =

$$\frac{0.0841 \times 0.29 + 0.0841 \times 0.29 + 0 \times 0.29 + 0 \times 0.29 + 0 \times 0 + 0.0232 \times 0 + 0 \times 0 + 0 \times 0.8321}{0.119 \times 0.29} \approx 0.634$$

Cosine Similarity (Q, D2) =

$$\frac{0.0841 \times 0.29 + 0 \times 0.29 + 0.0841 \times 0.29 + 0 \times 0.29 + 0 \times 0 + 0.0232 \times 0 + 0.08729 \times 0.29 + 0 \times 0}{0.119 \times 0.29} \approx 0.634$$

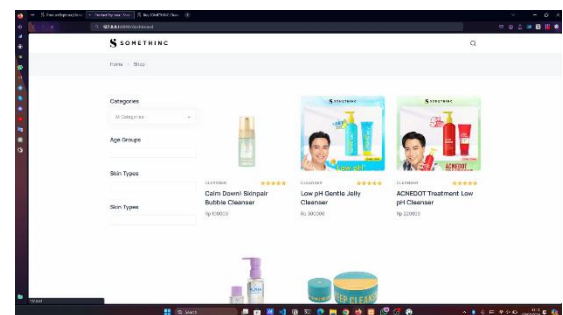
Tabel 7. Rangkng Hasil Rekomendasi

Term	Nilai Cosine	Rangking
D4	0.722	1
D7	0.722	2
D1	0.634	3
D2	0.634	4
D3	0.634	5
D6	0.634	6
D5	0.604	7
D8	0	8

3.4 Prototipe

a. Halaman Utama

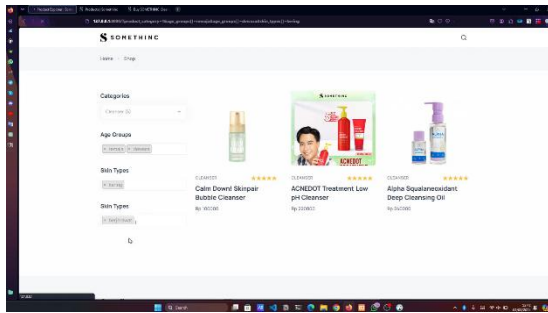
Di halaman beranda, ada opsi pemfilteran yang memungkinkan pengguna mempersempit pencarian berdasarkan kategori, usia, dan jenis kulit.



Gambar 2. Tampilan Halaman Utama

b. Tampilan Halaman Filter

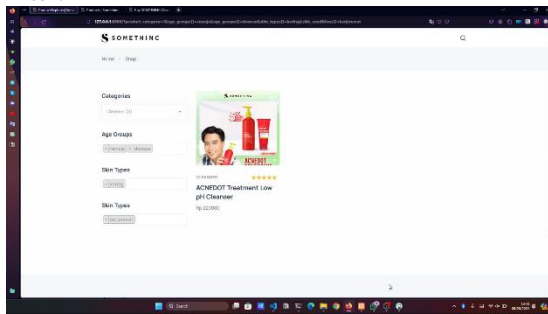
Menu pemfilteran digunakan ketika pengguna mengisi kriteria pemfilteran. Fungsi dari menu ini adalah untuk membantu pengguna menemukan rekomendasi yang sesuai dengan jenis kulit mereka dengan lebih mudah.



Gambar 3. Tampilan Halaman Filter

c. Tampilan Hasil Rekomendasi

Pada menu hasil rekomendasi dapat diakses setelah pengguna mengisi kriteria pada menu filter.



Gambar 4. Tampilan Halaman Rekomendasi

SIMPULAN

Hasil penelitian menunjukkan bahwa dokumen D4 dan D7 memiliki kesamaan tertinggi dengan dokumen pertanyaan, dengan nilai *cosine similarity* sebesar 0.722. Dokumen D1, D2, D3, dan D6 memiliki nilai *cosine similarity* yang sama, yaitu sebesar 0.634. Dokumen D5 memiliki nilai *cosine similarity* sebesar 0.604, sedangkan dokumen D8 memiliki nilai *cosine similarity* terendah, yaitu 0. Oleh karena itu, dokumen D4 dengan nama produk *Bee Power Propolis Manuka Honey Sleeping Mask* dan D7 dengan nama produk *PEPTINOL Granactive Retinoid Peptide Night Moisturizer Creme* adalah rekomendasi terbaik karena memiliki nilai kemiripan tertinggi.

Berdasarkan hasil dan analisis yang dilakukan, dapat disimpulkan bahwa sistem rekomendasi produk Somethinc berhasil dibangun dengan mengintegrasikan metode *content-based filtering*.

Untuk penelitian berikutnya, disarankan agar jumlah produk yang tersedia dalam sistem rekomendasi ditingkatkan agar mendukung hasil rekomendasi yang lebih akurat dan beragam. Selain itu, sistem ini juga dapat dikembangkan menjadi sebuah platform *mobile* berbasis Android.

UCAPAN TERIMA KASIH

Penulis ingin menyampaikan rasa terima kasih yang mendalam kepada Bapak Anief Fauzan Rozi atas arahan, bimbingan yang

diberikan selama proses penelitian ini. Kontribusinya sangat berarti dalam memandu penelitian ini dari awal hingga akhir. Semoga penelitian ini memberikan manfaat yang besar dan berkontribusi positif dalam pengembangan teknologi.

DAFTAR PUSTAKA

- [1] V. N. Yefta and D. Y. Bernanda, "Perancangan Sistem Informasi menggunakan TOGAF Dan Analisis Ward & Peppard pada SMA Santo Leo," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 6, no. 2, pp. 314–323, Apr. 2024, doi: 10.47233/jteksis.v6i2.1171.
- [2] K. Wakhidah, B. Budiman, and W. Winarti, "Rancangan Sistem Informasi Perpustakaan Berbasis Website Menggunakan Barcode Di Sekolah MA Raden Rahmat," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 5, no. 1, pp. 61–68, Feb. 2023, doi: 10.47233/jteksis.v5i1.740.
- [3] E. Ayuningrum, Y. Azhar, and G. I. Marthasari, "Sistem Rekomendasi Produk Skincare Korea Berbasis Web Menggunakan Metode Collaborative Filtering," *REPOSITOR*, vol. 4, no. 4, pp. 497–506, 2022.
- [4] M. I. Rahayu, M. Kasegrina Siregar, M. Desnia, and S. Bandung, "SISTEM REKOMENDASI PEMILIHAN PRODUK BASIC SKINCARE BERDASARKAN JENIS KULIT MENGGUNAKAN ALGORITMA SIMPLE ADDITIVE WEIGHTING (SAW)," 2022.
- [5] A. Risetdikti and J. Komunikasi, "'Somethinc' Pada Pengguna Instagram," vol. 13, no. 2, 2022, doi: 10.31294/jkom.
- [6] D. Ayu, N. Safitri, R. Helilintar, and L. S. Wahyuniar, "Sistem Rekomendasi Skincare Menggunakan Metode Content-Based Filtering dan Algoritma Apriori."
- [7] M. Alkaff, H. Khatimi, and A. Eriadi, "Sistem Rekomendasi Buku pada Perpustakaan Daerah Provinsi Kalimantan Selatan Menggunakan Metode Content-Based Filtering," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 1, pp. 193–202, Sep. 2020, doi: 10.30812/matrik.v20i1.617.
- [8] A. E. Wijaya and D. Alfian, "SISTEM REKOMENDASI LAPTOP MENGGUNAKAN COLLABORATIVE FILTERING DAN CONTENT-BASED FILTERING," *Jurnal Computech & Bisnis*, vol. 12, no. 1, pp. 11–27, 2018.
- [9] C. Fiarni, H. Maharani, and N. Calista, "Product Recommendation System Design Using Cosine Similarity and Content-based Filtering Methods," 2019.
- [10] S. Pumomo and F. A. Alijojo, "Sistem Peminjaman Barang Menggunakan QR Code Berbasis Aplikasi Android," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 6, no. 2, pp. 322–328, Apr. 2024, doi: 10.47233/jteksis.v6i2.1350.
- [11] M. Nurjannah, I. Fitri Astuti, and D. Program Studi, "PENERAPAN ALGORITMA TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK TEXT MINING Mahasiswa S1 Program Studi Ilmu Komputer FMIPA Universitas Mulawarman 2,3)," 2013.
- [12] R. Al Rasyid, D. Handayani, and U. Ningsih, "Penerapan Algoritma TF-IDF dan Cosine Similarity untuk Query Pencarian Pada Dataset Destinasi Wisata," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 8, no. 1, p. 2024, 2024, doi: 10.35870/jti.
- [13] B. U. Tri Wahyo and A. Widya Anggriawan, "Sekolah Tinggi Manajemen Informatika dan Komputer ASIA Malang 6 SISTEM REKOMENDASI PAKET WISATA SE-MALANG RAYA MENGGUNAKAN

- METODE HYBRID CONTENT BASED DAN COLLABORATIVE,” 2015.
- [14] M. Alkaff, H. Khatimi, and A. Eriadi, “Sistem Rekomendasi Buku pada Perpustakaan Daerah Provinsi Kalimantan Selatan Menggunakan Metode Content-Based Filtering,” *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 1, pp. 193–202, Sep. 2020, doi: 10.30812/matrik.v20i1.617.
- [15] L. E. Sari and W. Hadikurniawati, *SISTEM PENDUKUNG KEPUTUSAN PENENTUAN SKINCARE UNTUK KULIT WAJAH MENGGUNAKAN METODE DECISION TREE (STUDI KASUS PADA KOSMETIK WARDAH)*.
- [16] A. Laksito and M. R. Saputra, “Content Based VGG16 Image Extraction Recommendation,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 3, pp. 370–375, Jun. 2022, doi: 10.29207/resti.v6i3.3909.
- [17] C. Koloman, R. Maulana, R. Dwi, Z. Putri, and W. A. Harahap, “Sistem Rekomendasi Pekerjaan di bidang IT Menggunakan Algoritma Content-Based Filtering,” *Journal of Creative Student Research (JCSR)*, vol. 1, no. 6, pp. 78–88, 2023, doi: 10.55606/jcsrpolitama.v1i6.2992.
- [18] M. Fajriansyah, P. P. Adikara, and A. W. Widodo, “Sistem Rekomendasi Film Menggunakan Content Based Filtering,” 2021.
- [19] E. Tjipta, W. Sekolah, T. Manajemen Informatika, D. Komputer, and A. Malang, “78 Sekolah Tinggi Manajemen Informatika dan Komputer ASIA Malang PERANCANGAN INFORMATION RETRIEVAL (IR) BERBASIS TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF) UNTUK PERINGKASAN TEKS TUGAS KHUSUS BERBAHASA INDONESIA,” 2013.
- [20] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, “Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity,” 2018.
- [21] M. A. Rofiqi, Abd. C. Fauzan, A. P. Agustin, and A. A. Saputra, “Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query,” *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 1, no. 2, pp. 58–64, Dec. 2019, doi: 10.28926/ilkomnika.v1i2.18.
- [22] I. Cholissodin *et al.*, “PEMERINGKATAN PENCARIAN PADA BUKU PEDOMAN AKADEMIK FILKOM UB MENUJU MERDEKA BELAJAR DAN FREE E-BOOK PEMBELAJARAN SEBAGAI PROTOTYPE LOCAL SMART MICRO SEARCH ENGINE MENGGUNAKAN ALGORITMA PAGERANK DAN TF-IDF,” doi: 10.25126/jtiik.202184384.
- [23] M. A. Rofiqi, Abd. C. Fauzan, A. P. Agustin, and A. A. Saputra, “Implementasi Term-Frequency Inverse Document Frequency (TF-IDF) Untuk Mencari Relevansi Dokumen Berdasarkan Query,” *ILKOMNIKA: Journal of Computer Science and Applied Informatics*, vol. 1, no. 2, pp. 58–64, Dec. 2019, doi: 10.28926/ilkomnika.v1i2.18.
- [24] S. Fatonah Larasati, U. Ria Safitri, L. Puji Rahayu, F. Ekonomi, and U. Boyolali Jl Pandanaran No, “PENGARUH KUALITAS PRODUK, PROMOSI DAN POTONGAN HARGA TERHADAP KEPUTUSAN PEMBELIAN PRODUK WARDAH KOSMETIK (STUDI KASUS PADA KEPUTUSAN PEMBELIAN PRODUK WARDAH KOSMETIK DI TOKO EVIAA COSMETIK KARTASURA),” *Jurnal Ilmu Manajemen dan Akuntansi*, vol. 9, no. 2, p. 2021, 2021.
- [25] D. Septiani and I. Isabela, “SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia ANALISIS TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) DALAM TEMU KEMBALI INFORMASI PADA DOKUMEN TEKS”.
- [26] S. K. Dirjen *et al.*, “Terakreditasi SINTA Peringkat 2 Implementasi Metode Content Based Filtering Pada Aplikasi Pencarian Taman Penitipan Anak,” *masa berlaku mulai*, vol. 1, no. 3, pp. 163–169, 2017.
- [27] A. Riyani, M. Zidny Naf’an #2, and A. Burhanuddin, “Penerapan Cosine Similarity dan Pembobotan TF-IDF untuk Mendeteksi Kemiripan Dokumen,” 2019.
- [28] A. E. Wijaya and D. Alfian, “SISTEM REKOMENDASI LAPTOP MENGGUNAKAN COLLABORATIVE FILTERING DAN CONTENT-BASED FILTERING,” *Jurnal Computech & Bisnis*, vol. 12, no. 1, pp. 11–27, 2018.