

Sistem Rekomendasi Merek Parfum Lokal Menggunakan Content-Based Filtering dan Cosine Similarity

Stefanus Ardi Prasetyo^{1*}, Adi Lukman Maulana², Alvifah Diyan Safitri³, Fauzi Bondan Prihananto⁴, Tomy Anugerah Islami⁵, Vihi Atina⁶

¹Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa
^{1*}230103237@mhs.udb.ac.id

²Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa
²230103212@mhs.udb.ac.id

³Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa
³230103215@mhs.udb.ac.id

⁴Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa
⁴230103224@mhs.udb.ac.id

⁵Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa
⁵230103239@mhs.udb.ac.id

⁶Teknik Informatika/Illmu Komputer
Universitas Duta Bangsa
⁶ vihi_atina@udb.ac.id

Abstrak— Pertumbuhan merek parfum lokal menghadirkan semakin banyak pilihan aroma, harga, dan karakter produk bagi konsumen. Namun, banyaknya informasi tersebut dapat menimbulkan *information overload* sehingga konsumen kesulitan menentukan parfum yang sesuai dengan preferensi aromanya. Penelitian ini merancang sistem rekomendasi merek parfum lokal menggunakan pendekatan *Content-Based Filtering*, yaitu rekomendasi disusun berdasarkan kemiripan karakteristik produk tanpa bergantung pada riwayat penilaian pengguna lain. Data deskripsi parfum, terutama informasi aroma atau *notes*, diproses melalui *case folding*, *tokenization*, *stopword removal*, dan *stemming* agar bentuk teks lebih seragam. Selanjutnya, metode *Term Frequency-Inverse Document Frequency (TF-IDF)* digunakan untuk mengubah deskripsi menjadi vektor numerik dan memberikan bobot pada kata kunci yang merepresentasikan karakter produk. Nilai kemiripan antarproduk dihitung menggunakan algoritma *Cosine Similarity* sehingga parfum dengan skor kesamaan tertinggi dapat ditampilkan sebagai rekomendasi. Hasil rancangan sistem menunjukkan bahwa rekomendasi terbaik selama uji coba dengan parameter user top note: grapefruit, mid note: fruity, base note: musk, gender: unisex, situasi: versatile menghasilkan rekomendasi terkuat, yaitu parfum Infinity dari merek Avila, dengan nilai *similarity* 70%. Setelah dilakukan evaluasi, sistem mendapat hasil 100% untuk *precision*, 71% untuk *recall*, dan menghasilkan *Fi-Score* sebesar 83%. Sistem dapat dikatakan cukup akurat dan diharapkan mampu menyajikan alternatif parfum lokal yang lebih relevan dengan minat pengguna, mempercepat proses pemilihan produk, dan mengurangi dampak *information overload*.

Kata kunci— Sistem Rekomendasi, Parfum Lokal, *Content-Based Filtering*, *TF-IDF*, *Cosine Similarity*.

Abstract—The rapid growth of local perfume brands offers consumers an expanding array of scent options, price points, and product characteristics. However, this abundance of information can lead to information overload, making it difficult for consumers to find a perfume that truly matches their fragrance preferences. To address this, the study designs a local perfume brand recommendation system using a *Content-Based Filtering* approach, in which recommendations are generated based on the similarity of product characteristics, independent of other users' rating histories. Perfume description data—specifically fragrance notes—undergoes preprocessing via *case folding*, *tokenization*, *stop-word removal*, and *stemming* to standardize the text format. The *Term Frequency-Inverse Document Frequency (TF-IDF)* method is then applied to convert these descriptions into numerical vectors and assign weights to keywords representing the product's characteristics. Product similarity is calculated using the *Cosine Similarity* algorithm, allowing perfumes with the highest similarity scores to be surfaced as recommendations. System testing demonstrates that using specific user parameters (top note: grapefruit, mid note: fruity, base note: musk, gender: unisex, occasion: versatile) yields the strongest recommendation for the perfume "Infinity" by the brand Avila, with a 70% similarity score. Following the evaluation, the system achieved 100% precision, 71% recall, and an *Fi-score* of 83%. The system proves to be highly reliable and is expected to effectively deliver relevant local perfume alternatives tailored to user interests, streamline the selection process, and mitigate the impact of *information overload*.

Keywords— Recommendation System, Local Perfume, *Content-Based Filtering*, *TF-IDF*, *Cosine Similarity*.

I. PENDAHULUAN

Industri kecantikan dan perawatan diri terus berkembang, termasuk pada segmen parfum lokal. Banyaknya merek dan varian aroma yang muncul memberi konsumen lebih banyak pilihan, tetapi pada

saat yang sama dapat menimbulkan kebingungan ketika pengguna harus menentukan produk yang paling sesuai dengan preferensinya. Kondisi tersebut berkaitan dengan *information overload*, yaitu situasi ketika jumlah informasi yang diterima melebihi

kemampuan seseorang untuk memilah informasi yang relevan [1]. Dalam konteks pemilihan parfum, kelebihan informasi dapat menurunkan kualitas keputusan karena konsumen harus membandingkan banyak atribut seperti karakter aroma, konsentrasi, ukuran, harga, dan segmentasi penggunaan.

Sistem rekomendasi dapat digunakan sebagai mekanisme penyaringan informasi karena mampu memperkirakan *item* yang berpotensi disukai pengguna [2]. Salah satu pendekatan yang banyak digunakan adalah *Content-Based Filtering* (CBF). Metode ini menghasilkan rekomendasi berdasarkan kemiripan atribut konten antar-item, sehingga tidak bergantung pada data interaksi atau penilaian pengguna lain [3]. CBF telah diterapkan pada beberapa domain, misalnya rekomendasi buku akademik, buku perpustakaan sekolah, film berbasis genre, dan film berdasarkan sinopsis [1]–[5]. Pada implementasi berbasis teks, metode *Term Frequency-Inverse Document Frequency* (TF-IDF) sering digunakan untuk mengekstraksi dan memberi bobot pada kata kunci penting, sedangkan *Cosine Similarity* digunakan untuk mengukur kedekatan antarvektor fitur [2], [4].

Fleksibilitas CBF, TF-IDF, dan *Cosine Similarity* juga terlihat pada penelitian di bidang lagu, kosmetik, destinasi wisata, resep masakan, serta lowongan pekerjaan [1], [6]–[10]. Namun, penerapan metode tersebut pada rekomendasi merek parfum lokal berbasis karakteristik aroma atau *notes* masih belum banyak dibahas. Berdasarkan kondisi tersebut, penelitian ini bertujuan merancang sistem rekomendasi merek parfum lokal dengan pendekatan *Content-Based Filtering*. Deskripsi produk diolah menggunakan TF-IDF untuk membentuk representasi fitur, kemudian tingkat kemiripan antarproduk dihitung menggunakan *Cosine Similarity*. Hasil rancangan diharapkan dapat membantu konsumen memperoleh rekomendasi parfum lokal yang lebih sesuai dengan preferensi aroma dan mengurangi dampak kelebihan informasi.

II. METODOLOGI PENELITIAN

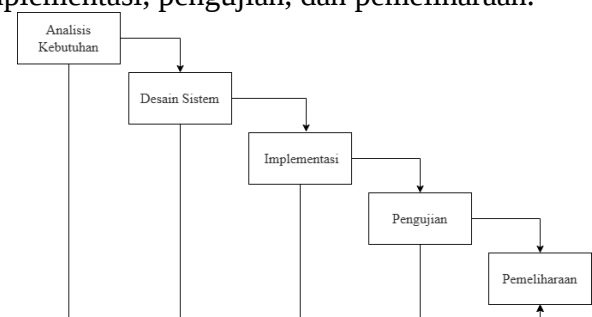
Metodologi penelitian ini menjelaskan tahapan komputasi dan analisis yang digunakan dalam perancangan sistem rekomendasi merek parfum lokal berbasis *Content-Based Filtering*. Secara umum, proses penelitian meliputi pengumpulan

data, pra-pemrosesan teks, pembobotan fitur, penghitungan kemiripan, serta evaluasi sistem.

Objek penelitian berupa data katalog parfum lokal yang memuat atribut tekstual dan atribut pendukung produk. Atribut aroma menjadi fokus utama karena merepresentasikan karakter parfum, sedangkan atribut lain seperti harga, ukuran, konsentrasi, situasi penggunaan, dan target pengguna digunakan sebagai informasi tambahan untuk memperkaya deskripsi produk. Bagian yang belum memiliki data lengkap tetap dapat dilewati sementara dan dilengkapi pada tahap pembaruan *dataset*.

1. Metode Pengembangan Sistem (Waterfall)

Pengembangan sistem pada penelitian ini menggunakan model *Waterfall*. Model ini memiliki alur kerja berurutan sehingga setiap tahap diselesaikan terlebih dahulu sebelum melanjutkan ke tahap berikutnya. Tahapan yang digunakan mencakup analisis kebutuhan, desain sistem, implementasi, pengujian, dan pemeliharaan.



Gambar 1. Waterfall Sistem

1. Analisis Kebutuhan (Requirement)

Tahap ini berfokus pada identifikasi permasalahan dan kebutuhan sistem. Pada penelitian ini, analisis dilakukan dengan memetakan masalah *information overload* dalam pemilihan parfum serta menentukan kebutuhan data berupa katalog dan deskripsi *notes* aroma dari merek parfum lokal.

2. Desain Sistem (Design)

Tahap desain digunakan untuk merancang struktur data, alur pemrosesan, dan antarmuka sistem. Informasi produk direncanakan dalam format data terstruktur, misalnya CSV, agar atribut parfum seperti nama produk, merek, harga, konsentrasi, ukuran, *notes*, situasi penggunaan, dan jenis kelamin dapat diproses secara konsisten.

3. Implementasi (Implementation)

Tahap implementasi menerjemahkan rancangan ke dalam program. Pada tahap ini, data deskripsi parfum diproses melalui *preprocessing*, kemudian bobot fitur dihitung menggunakan *TF-IDF* dan kemiripan antarproduk dihitung dengan *Cosine Similarity*.

4. Pengujian (Verification/Evaluasi)

Tahap pengujian dilakukan untuk memastikan sistem berjalan sesuai kebutuhan. Pengujian fungsional dapat dilakukan menggunakan *Black Box Testing*, sedangkan kualitas hasil rekomendasi dapat dievaluasi menggunakan metrik *Precision*, *Recall*, dan *F1-Score*.

5. Pemeliharaan (Maintenance)

Tahap pemeliharaan mencakup perbaikan jika ditemukan kendala setelah sistem digunakan serta pembaruan data ketika terdapat produk parfum lokal baru yang perlu dimasukkan ke dalam basis data.

III. HASIL DAN PEMBAHASAN

A. Analisis Kebutuhan

1. Pengumpulan Data

Data yang dibutuhkan dalam penelitian ini berupa informasi produk dan deskripsi merek parfum lokal yang didapat dengan melakukan pencarian di internet dan di website pengumpul data seperti *Kaggle*. Berikut sampel tabel dari dataset yang akan digunakan dengan total jumlah data 35 produk.

Tabel 1. Kumpulan Sampel Data

perfume	brand	top notes	mid notes	base notes	situasi	gender
Es Teh Panda	Semerbak	lavender, white tea, mint	pandan, rose, green tea	sandalwood, patchouli	Versatile	Unisex
Orion	Velixir	Bergamot, Cinnamon	Bourbon, Vanilla, Praline, Ambroxan	Woody Notes, Amber, Musk, Heliotropine	Versatile	Unisex
Deep Dive	Oullu	lemon blossom, barley, ozonic	sage, hinoki	oak moss, cashmere wood	Versatile	Unisex
Infinity	Avila	lemon, lavender, grapefruit	cardamom, coriander, jasmine	sandal, musk, Patchouli, moss	Versatile	Unisex

		ozone, mandarin, bergamot	sambac, fruity, pink pepper, caramel	amber, vanilla		
Beloved	Permenent	Mandarin, Bergamot, Ozone, Blackcurrant, Gardenia	Jasmine, Neroli, Rose, Muguet, Prune, Lactone	Musk, Sandal, Amber, Vanilla, Moss, Benzoin	Versatile	Female
Voeu Provence	Verites	Mandarin, Pear, Blackcurrant, Pink Pepper	Jasmine, Violet, Tuberoses, Ylang-Ylang, Watery	Sandalwood, Caramel, Powdery, Amberwood, Musks	Versatile	Unisex
Atheia the Smilax Ornata	GHAIER	Chinotto, Sarsaparilla	Ylang-Ylang, Neroli, Geranium, Carnation	Labdanum, Vanilla, Birch, Star anise	Versatile	Unisex
Madeleine	Azwarly Perfume	Red Pepper, Nutmeg, Plum, Apple	Freesia, Peony, Jasmine	Musk, Amber, Sandalwood	Versatile	Unisex
Fields of Ubud	Project 1945	Ambrette, Seeds, Buchu Leaves, Bergamot, Elemi	Lavender, Sea Salt, Sage, Dried Fruit	Vetiver, Java, Sandalwood, Oakmoss	Day	Unisex

Atribut yang digunakan meliputi id parfum, nama parfum, merek parfum, harga, ukuran, jenis konsentrat, *top notes*, *middle notes*, *base notes*, situasi penggunaan, dan jenis kelamin target. Atribut tersebut dipilih karena dapat merepresentasikan karakteristik produk yang diperlukan dalam proses rekomendasi.

Atribut *top notes*, *middle notes*, dan *base notes* merupakan fitur utama karena ketiganya menjelaskan komposisi aroma dari fase awal hingga akhir pemakaian. Selain itu, pemilihan gender dan situasi menjadi fitur sekunder untuk menentukan rekomendasi. Fitur tersebut digunakan sebagai korpus yang akan diproses pada tahap *preprocessing* dan pembobotan *TF-IDF*.

Validasi data dilakukan dengan memastikan setiap produk memiliki identitas yang unik dan deskripsi aroma yang dapat diproses. Jika terdapat nilai kosong pada atribut pendukung, data tersebut dapat

tetap dipakai selama atribut *notes* utama tersedia. Sebaliknya, produk dengan deskripsi aroma yang kosong perlu ditunda atau dilengkapi terlebih dahulu agar tidak menurunkan kualitas rekomendasi.

B. Desain Sistem

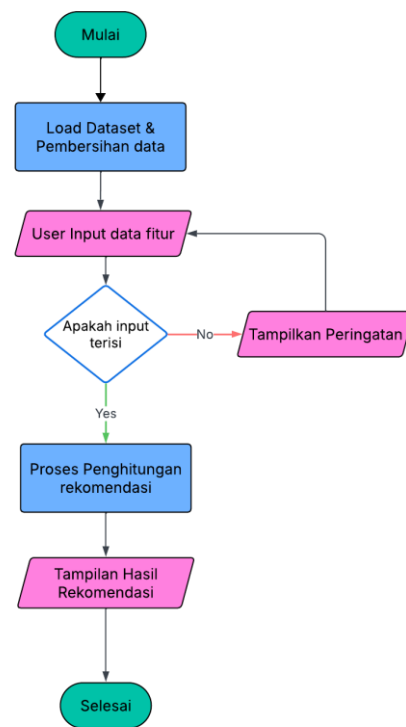
Sistem akan menggunakan Python sebagai bahasa pemrograman yang akan dijalankan dengan Streamlit secara sederhana. Berikut beberapa tahapan dalam desain sistem

1. Pra-Pemrosesan Data Teks (Text Preprocessing)

- Pembersihan teks (*case folding/cleaning*): Tahap ini menghapus angka, tanda baca, dan karakter khusus, serta mengubah seluruh huruf menjadi huruf kecil agar format teks konsisten.
- Tokenisasi (*tokenization*): Teks deskripsi parfum dipecah menjadi token atau kata tunggal sehingga setiap unsur aroma dapat dianalisis secara terpisah.
- Stopword removal*: Kata umum seperti “dan”, “atau”, serta kata penghubung lain yang tidak membedakan karakteristik aroma dihapus agar fitur yang tersisa lebih informatif.
- Stemming*: Kata turunan dikembalikan ke bentuk dasar agar istilah dengan makna serupa dapat dikenali secara konsisten.

2. Alur Kerja Sistem

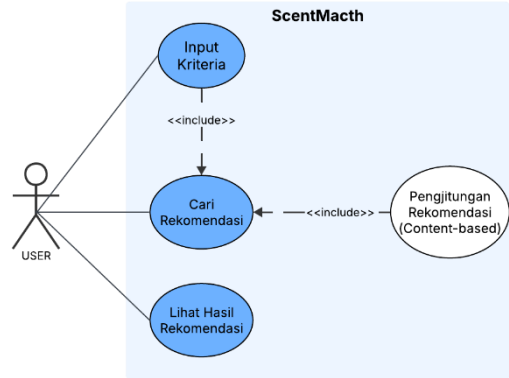
a) Flowchart



Gambar 2. Flowchart Sistem

Alur sistem rekomendasi dimulai dengan memuat dataset katalog parfum, yang dilanjutkan dengan tahap pembersihan data teks seperti case folding, tokenization, stopwords removal, dan stemming. Setelah data siap, pengguna diminta untuk memasukkan fitur atau kriteria parfum acuan sebagai input sistem. Sistem kemudian akan memvalidasi input tersebut; jika kosong, maka akan muncul peringatan, namun jika terisi, maka proses berlanjut ke tahap komputasi. Pada tahap penghitungan, sistem menerapkan pendekatan Content-Based Filtering dengan mengubah deskripsi teks menjadi vektor berbobot menggunakan TF-IDF dan menghitung jarak kedekatan aromanya melalui algoritma Cosine Similarity. Sebagai hasil akhir, sistem menampilkan daftar rekomendasi parfum lokal yang memiliki tingkat kemiripan paling tinggi dengan pilihan pengguna sebelum akhirnya proses dinyatakan selesai.

b) Use Case



Gambar 3. Use Case

Diagram use case bernama ScentMacth ini mengilustrasikan interaksi fungsional antara pengguna (user) dengan sistem rekomendasi parfum lokal. Interaksi pertama menunjukkan bahwa pengguna dapat menginput kriteria yang berisi pilihan parfum acuan sebagai dasar pencarian utama. Proses input tersebut saling terhubung (termasuk) dengan tahapan pencarian rekomendasi, di mana sistem akan merespons dan memproses permintaan berdasarkan masukan dari pengguna. Selama proses pencarian berlangsung, sistem secara internal menjalankan penghitungan rekomendasi menggunakan pendekatan Content-Based Filtering untuk mengukur tingkat kemiripan karakteristik aromanya. Terakhir, pengguna dapat berinteraksi dengan sistem untuk melihat hasil rekomendasi yang menampilkan daftar alternatif parfum lokal dengan urutan nilai kemiripan tertinggi.

3. Input dan Output Sistem

Input utama sistem berupa data parfum yang telah tersimpan dalam *dataset* dan pilihan parfum acuan dari pengguna. Output sistem berupa daftar produk parfum lokal yang diurutkan berdasarkan nilai kemiripan tertinggi. Setiap hasil rekomendasi dapat dilengkapi dengan informasi nama parfum, merek, *notes* aroma, dan skor kemiripan agar pengguna dapat memahami alasan rekomendasi yang diberikan.

C. Implementasi

1. Ekstraksi Fitur dan Pembobotan TF-IDF

Deskripsi parfum yang telah melalui tahap *preprocessing* dikonversi ke bentuk vektor numerik

menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)*. Metode ini menilai pentingnya suatu kata dalam satu dokumen berdasarkan frekuensi kemunculannya pada dokumen tersebut dan kelangkaannya pada keseluruhan koleksi data. Kata yang spesifik seperti musk, citrus, vanilla, woody, atau floral akan memperoleh bobot lebih tinggi ketika kata tersebut mampu membedakan satu produk dari produk lain.

Hasil pembobotan *TF-IDF* berupa matriks fitur dengan baris yang merepresentasikan produk parfum dan kolom yang merepresentasikan istilah hasil *preprocessing*. Matriks ini menjadi dasar perhitungan kemiripan karena setiap parfum sudah memiliki representasi numerik yang dapat dibandingkan secara sistematis. Untuk penghitungan TF, sistem menggunakan penghitungan mentah (*Raw Count*) untuk menghitung kemunculan sebuah kata secara absolut. Untuk penghitungan IDF, gunakan rumus berikut.

$$IDF = \left(\frac{N}{df}\right) + 1$$

(N) Adalah jumlah total seluruh parfum dalam data, (df) merupakan jumlah parfum yang mengandung parameter yang dimaksud. Penambahan angka 1 agar aroma yang muncul di *semua* dokumen tidak memiliki nilai IDF yang sama dengan 0. Dengan begitu, aroma universal tersebut tetap memiliki bobot minimal dan tidak hilang dalam perhitungan. Lalu perhitungan TF-IDF menggunakan rumus berikut.

$$TF-IDF(t, d) = TF(t, d) \times IDF(t)$$

Pada rumus tersebut, TF menunjukkan frekuensi istilah t dalam dokumen d, sedangkan IDF menunjukkan tingkat kelangkaan istilah pada seluruh dokumen. Semakin spesifik suatu istilah terhadap satu parfum, semakin besar kontribusi istilah tersebut dalam membentuk profil produk. Berikut contoh tf-idf dari parameter user dengan data total jumlah parfum dalam dataset 35 produk dan total *unique vocabulary* pada notenya adalah 166 kata.

Tabel 2. Perhitungan TF-IDF

No	term	tf	df	idf	tf-idf
1	fruity	1	1	2,5441	2,5441
2	grapefruit	1	4	1,942	1,942
3	musk	1	14	1,3979	1,3979

Untuk gender dan situasi akan dihitung terpisah karena menjadi penentu sekunder.

2. Perhitungan Kemiripan Menggunakan Cosine Similarity

Setelah vektor *TF-IDF* terbentuk, sistem menghitung kedekatan antarproduk menggunakan *Cosine Similarity*. Nilai kemiripan berada pada rentang 0 sampai 1. Semakin mendekati 1, semakin tinggi kesamaan karakteristik aroma antarproduk. Dengan cara ini, sistem dapat menampilkan daftar parfum lokal yang memiliki karakter aroma paling dekat dengan produk yang diminati atau dipilih oleh pengguna.

$$Cosine(A,B) = \frac{(A \times B)}{(|A| \times |B|)}$$

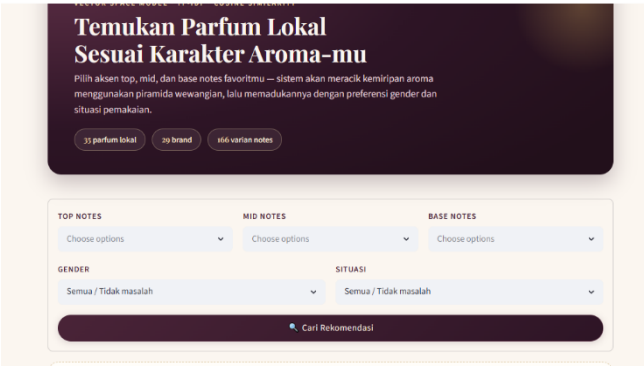
Pada proses rekomendasi, produk yang dipilih pengguna dijadikan *item* acuan. Sistem kemudian menghitung nilai kemiripan *item* acuan terhadap seluruh produk lain dalam *dataset*, mengurutkan hasil dari skor tertinggi, dan menampilkan sejumlah rekomendasi teratas. Produk acuan dikeluarkan dari daftar hasil agar sistem tidak merekomendasikan *item* yang sama.

Jika dua produk memiliki komposisi kata kunci aroma yang mirip, sudut antarvektor keduanya akan semakin kecil dan nilai *Cosine Similarity* semakin tinggi. Nilai ini kemudian digunakan sebagai dasar pemeringkatan produk dalam daftar rekomendasi. Berikut hasil similarity untuk hasil penghitungan diatas 50%.

Tabel 3. Hasil Perhitungan Cosine Similarity

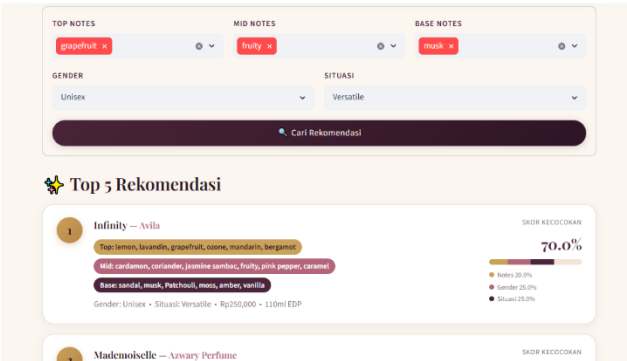
No	perfume	brand	final_score
1	Infinity	Avila	0,699
2	Mademoiselle	Azwary Perfume	0,543
3	Orion	Velixir	0,544
4	Deep Dive	Oullu	0,5
5	Es Teh Pandan	Semberak	0,5
6	Atheia the Smilax Ornata	GHAIER	0,5
7	Voieu Provence	Verites	0,5
8	Beloved	Permenent	0,283
9	Fields of Ubud	Project 1945	0,25

3. Tampilan Sistem
a) Tampilan Dashboard



Gambar 3. Dashboard awal

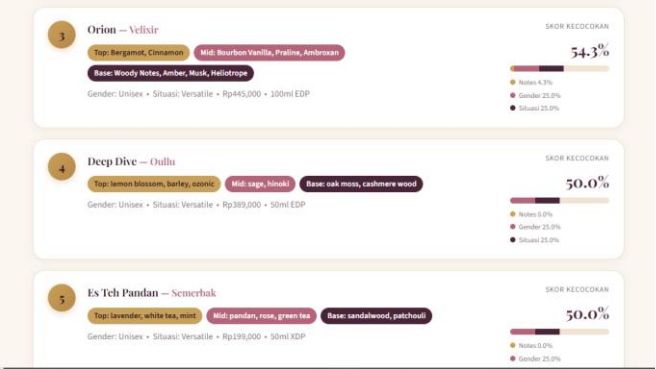
Tampilan ini merupakan antarmuka utama pengguna (Dashboard) di mana pengguna dapat mencari parfum lokal sesuai dengan karakter aroma yang diinginkan. Pada halaman ini, terdapat beberapa opsi input bagi pengguna untuk memilih preferensi piramida wewangian, yang terdiri dari parameter Top Notes, Mid Notes, dan Base Notes. Selain parameter aroma, pengguna juga dapat menentukan preferensi tambahan untuk spesifikasi Gender dan Situasi pemakaian. Setelah kriteria dipilih, pengguna dapat mengeklik tombol "Cari Rekomendasi" agar sistem memproses data tersebut.



Gambar 4. Hasil rekomendasi 1

Gambar 4 menunjukkan antarmuka hasil komputasi sistem yang menampilkan bagian atas dari daftar "Top 5 Rekomendasi". Pada urutan ini, sistem menampilkan produk parfum dengan tingkat kemiripan tertinggi terhadap kriteria pengguna, yaitu "Infinity — Avila" di peringkat pertama dan "Mademoiselle" di peringkat kedua. Setiap hasil yang disajikan memuat informasi detail produk seperti nama parfum, merek, rincian notes (Top,

Mid, Base), kesesuaian gender dan situasi, harga, volume, serta indikator ”Skor Kecocokan” (dalam persentase).



Gambar 5. Hasil rekomendasi 2

Tampilan ini merupakan kelanjutan dari daftar rekomendasi sebelumnya, yang menampilkan sisa dari lima rekomendasi teratas. Pada gambar ini, terlihat produk parfum yang menduduki peringkat ketiga hingga kelima, seperti ”Orion — Velixir”, ”Deep Dive — Oullu”, dan ”Es Teh Pandan — Semerbak”. Seluruh produk yang direkomendasikan pada bagian ini juga dilengkapi dengan rincian parameter notes, kelengkapan informasi produk, dan persentase ”Skor Kecocokan” yang memvisualisasikan seberapa dekat karakteristik produk tersebut dengan preferensi pengguna.

D. Pengujian (Verification/Evaluasi)

1. Evaluasi Sistem

Evaluasi sistem dilakukan melalui dua pendekatan. Pertama, pengujian fungsional menggunakan *Black Box Testing* untuk memastikan setiap fitur, alur *input*, dan keluaran sistem berjalan sesuai rancangan. Kedua, evaluasi performa rekomendasi dilakukan secara kuantitatif menggunakan *Precision*, *Recall*, dan *Fi-Score*. *Precision* digunakan untuk melihat ketepatan rekomendasi yang ditampilkan, *Recall* mengukur kemampuan sistem menemukan *item* relevan, sedangkan *Fi-Score* memberikan gambaran keseimbangan antara *Precision* dan *Recall*. Berikut untuk ahsil blackbox testing.

Tabel 4. Hasil Keberhasilan Sistem

Fitur / fungsi sistem	Hasil uji
-----------------------	-----------

Halaman utama	Berhasil
Pemilihan parameter	Berhasil
Top Notes	
Pemilihan parameter	Berhasil
Mid Notes	
Pemilihan parameter	Berhasil
Base Notes	
Pemilihan parameter	Berhasil
Gender	
Pemilihan parameter	Berhasil
Situasi	
Tombol fungsi rekomendasi	Berhasil
Tampilan hasil rekomendasi	Berhasil

Metrik *Precision*, *Recall*, dan *Fi-Score* dapat dihitung apabila tersedia label relevansi atau penilaian pakar terhadap daftar rekomendasi. Berikut hasil salah satu uji coba rekomendasi, dengan user parameter top note: graperuit, mid note: fruity, base note: musk, gender: unisex, situasi: versatile.

Tabel 5. Hasil Uji Evaluasi Sistem

No	Nama	Parameter	Nilai
1	Infinity/ Avila	Top: lemon, lavender, grapefruit, ozone, mandarin, bergamot, Mid: cardamon, dll, Base: sandal, musk, Patchouli, moss, Gender: Unisex, Situasi: Versatile amber, vanilla	70%
2	Mademoiselle/ Azwary Perfume	Top: Red Pepper, Nutmeg, Plum, Apple, Mid: Freesia, Peony, Jasmine, Base: Musk, Amber, Sandalwood, Gender: Unisex, Situasi: Versatile	54,4%

3	Orion/ Velixir	Top: lemon blossom, barley, ozonic, Mid: sage, hinoki, Base: oak moss, cashmere wood, Gender: Unisex, Situasi: Versatile	54,3%
4	Deep Dive/Oullu	Top: Bergamot, Cinnamon, Mid: Bourbon Vanilla, Praline, Ambroxan, Base: Woody Notes, Amber, Musk, Heliotrope, Gender: Unisex, Situasi: Versatile	50%
5	Teh Pandan/ Semerbak	Top: lavender, white tea, mint, Mid: pandan, rose, green tea, Base: sandalwood, patchouli, Gender: Unisex, Situasi: Versatile	50%

Dari hasil tersebut maka didapatkan hasil evaluasi sebagai berikut:

1. *Precision*

$$Precision = \frac{\text{Data relevan tertampil}}{5} Precision = \frac{5}{5} = 100\%$$

2. *Recall*

$$Recall = \frac{\text{Data relevan tertampil}}{\text{Total data relevan}} Recall = \frac{5}{7} = 71\%$$

3. *FI-Score*

$$Fi\ Score = 2 \times \frac{Precision \times Recall}{(Precision + Recall)}$$

$$Fi\ Score = 2 \times \frac{100 \times 71}{(100 + 71)} = 83\%$$

Precision yang tinggi menunjukkan bahwa sebagian besar *item* yang direkomendasikan

memang relevan. *Recall* yang memiliki nilai 71% dikatakan cukup untuk menunjukkan bahwa data cukup lengkap, tetapi masih ada data yang tidak ditampilkan karena desain aplikasi hanya menampilkan 5 teratas, sedangkan total data relevan diambil saat produk memiliki nilai similarity 50% ke atas, sehingga data yang cukup kuat untuk direkomendasikan tidak ditampilkan. *Fi-Score* mendapat 83%, artinya sistem cukup akurat untuk merekomendasikan parfum yang diinginkan user, namun sistem harus bisa ditingkatkan lagi agar data yang ditampilkan lebih lengkap.

E. Pemeliharaan

Tahap pemeliharaan dilakukan untuk memastikan sistem rekomendasi tetap berjalan stabil dan relevan setelah digunakan. Kegiatan pemeliharaan meliputi perbaikan *bug*, pengecekan fungsi pencarian rekomendasi, pembaruan *dataset* parfum lokal, serta validasi ulang atribut aroma agar data yang digunakan tetap konsisten.

Pemeliharaan juga mencakup penyesuaian sistem terhadap perubahan kebutuhan pengguna. Apabila terdapat merek atau varian parfum baru, data produk dapat ditambahkan ke dalam basis data dan diproses kembali melalui tahapan *preprocessing*, pembobotan *TF-IDF*, serta perhitungan *Cosine Similarity*. Dengan demikian, sistem dapat terus menghasilkan rekomendasi yang relevan dan mengikuti perkembangan produk parfum lokal.

IV. KESIMPULAN

Penelitian ini menghasilkan rancangan sistem rekomendasi merek parfum lokal berbasis *Content-Based Filtering* dengan memanfaatkan *TF-IDF* untuk pembobotan fitur teks dan *Cosine Similarity* untuk menghitung kemiripan antarproduk. Tahapan penelitian dimulai dari pengumpulan data, pra-pemrosesan teks, pembentukan vektor fitur, perhitungan kemiripan, evaluasi, hingga pemeliharaan sistem. Rancangan ini menunjukkan bahwa karakteristik aroma parfum dapat diolah sebagai dasar rekomendasi sehingga pengguna memperoleh daftar produk yang lebih sesuai dengan preferensi aroma. Dalam hasil pengujian, sistem dapat merekomendasikan produk parfum Infinity dari merek Avila sebagai hasil rekomendasi terkuat dengan skor similarity 70% dengan parameter yang

dimasukkan user berupa grapefruit sebagai top note, fruity sebagai mid note, musk sebagai base note, dengan parameter gender unisex, dan situasi versatile. Sistem dapat dikatakan cukup akurat dengan evaluasi fi-score sebesar 83 % dan mampu merekomendasikan produk sesuai preferensi pengguna dengan baik, namun masih perlu penyempurnaan lebih lanjut.

Tujuan penelitian untuk membantu konsumen menghadapi banyaknya pilihan parfum lokal dapat tercapai melalui mekanisme penyaringan produk berdasarkan kesamaan konten. Sistem ini berpotensi mengurangi *information overload* dan mempercepat proses pengambilan keputusan. Pengembangan lanjutan dapat dilakukan dengan memperluas *dataset*, menambahkan umpan balik pengguna, serta menambahkan algoritma baru agar rekomendasi lebih akurat.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Duta Bangsa Surakarta, khususnya Fakultas Ilmu Komputer, atas dukungan dan fasilitas yang diberikan selama penyusunan penelitian ini. Penulis juga menyampaikan apresiasi kepada dosen pembimbing, rekan satu tim, serta semua pihak yang telah membantu dalam proses pengumpulan data, penyusunan naskah, dan pengembangan rancangan sistem rekomendasi.

REFERENSI

- [1] Y. S. Anggraeni, S. D. Saputri, T. Azzahra, A. G. Verrel, and I. Artikel, "Jurnal Informatika: Jurnal pengembangan IT Implementasi Content-Based Filtering Menggunakan TF-IDF dan Cosine Similarity untuk Rekomendasi Buku Akademik Mahasiswa," vol. 11, no. 2, pp. 213–225, 2026, doi: 10.30591/jpit.v11i2.10055.
- [2] T. Wahyu, I. Permadani, D. A. Prasetya, E. Daniati, and T. S. Informasi, "Sistem Rekomendasi Film Berdasarkan Genre Menggunakan Metode Content-Based Filtering dengan Algoritma Cosine Similarity," vol. 9, no. 1995, pp. 1835–1843.
- [3] D. Velamentosa and E. Zuliarso, "Sistem Rekomendasi Film Menggunakan Metode Content-Based Filtering," vol. 9, no. 2, pp. 2918–2922, 2025.
- [4] R. Ardiansyah et al., "Jurnal Computer Science and Information Technology (CoSciTech) Sistem Rekomendasi Buku Perpustakaan Sekolah Menggunakan Metode Content-Based Filtering," vol. 4, no. 2.
- [5] J. Maheswara, A. Saputra, and L. M. Huizen, "Sistem Rekomendasi Film pada Platform Streaming Menggunakan Metode Content-Based Filtering," vol. 15, no. 1, pp. 10–21, 2024.
- [6] A. Y. Timur and A. N. Rohman, "Indonesia Metode Content-Based Filtering dan Cosine Similarity," vol. 13, no. 1, pp. 1415–1423, 2025.
- [7] C. Koloman, R. Maulana, R. Dwi, Z. Putri, and W. Abadi, "Sistem Rekomendasi Pekerjaan di Bidang IT Menggunakan Algoritma Content-Based Filtering," vol. 1, no. 6, 2023.
- [8] W. Di and A. Tengah, "Implementasi Content-Based Filtering dengan TF-IDF dan Cosine Similarity untuk Sistem Rekomendasi Destinasi," vol. 10, no. 2, pp. 1329–1339, 2025.
- [9] N. O. Idris and F. Pontooyo, "Sistem Rekomendasi Produk Makeup Berbasis Content-Based Filtering dengan TF-IDF dan Cosine Similarity," KETIK: Jurnal Informatika, vol. 2, no. 06, pp. 24–32, 2025, doi: 10.70404/ketik.v2i06.311.
- [10] N. N. K. Sari, R. Priskila, and P. B. A. A. Putra, "Implementasi Content-Based Filtering Menggunakan TF-IDF and Cosine Similarity untuk Sistem Rekomendasi Resep Masakan," Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika, 2024.