

# Sistem Rekomendasi Parfum Lokal Berbasis Web Menggunakan Content-Based Filtering Dan Cosine Similarity

Mochamad Wahyu Sultan Fatahila<sup>1\*</sup>, Tanhella Zein Vitadiar<sup>2</sup>

<sup>1,2</sup>Teknik Informatika, Universitas Hasyim Asy'ari, Indonesia

<sup>1\*</sup>wahyufatahila@mhs.unhasy.ac.id, <sup>2</sup>tanhellavitadiar@unhasy.ac.id

**Abstrak:** Industri parfum lokal saat ini mengalami pertumbuhan yang pesat, namun pada praktiknya hal ini memicu fenomena *information overload* bagi konsumen yang sering kesulitan membedakan varian aroma karena proses pemilihan di platform digital masih bergantung pada deskripsi tekstual yang kurang interaktif dan belum menyediakan pencocokan preferensi secara personal. Penelitian ini bertujuan mengembangkan sistem rekomendasi parfum lokal berbasis web yang mengintegrasikan metode *Content-Based Filtering* dengan teknik pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) dan algoritma *Cosine Similarity* untuk menilai kemiripan kriteria pengguna terhadap target produk dalam basis data. Metode penelitian menggunakan pendekatan pengembangan perangkat lunak yang mencakup identifikasi masalah, studi literatur, pengumpulan data, perancangan sistem, implementasi, dan pengujian. Perancangan sistem dilakukan menggunakan *flowchart*, *use case diagram*, dan *entity relationship diagram*. Sistem memproses masukan kriteria dari pengguna meliputi aroma, *occasion*, gender, harga, dan konsentrasi, mengekstraksinya menjadi vektor bobot, lalu menghitung nilai jarak (*similarity*) untuk menghasilkan klasifikasi peringkat produk. Hasil pengujian menunjukkan bahwa aplikasi mampu menjalankan fungsi utama pencarian, menampilkan katalog, menerima input preferensi, dan memberikan umpan balik rekomendasi otomatis secara akurat. Pengujian presisi terhadap 20 sampel skenario pencarian menunjukkan 17 data relevan (*True Positive*) dan 3 data tidak relevan (*False Positive*), dengan tingkat akurasi sistem sebesar 85%. Temuan ini menunjukkan bahwa kombinasi TF-IDF dan *Cosine Similarity* layak diterapkan sebagai mesin rekomendasi interaktif untuk membantu konsumen memilih parfum secara lebih mandiri, terukur, dan akurat, serta meminimalisir risiko kesalahan pembelian (*blind buy*).

**Kata Kunci:** parfum lokal; web; content-based filtering; TF-IDF; cosine similarity; sistem rekomendasi

**Abstract:** The local perfume industry is currently experiencing rapid growth, yet in practice, this triggers an information overload phenomenon for consumers who often find it difficult to differentiate aroma variants because the selection process on digital platforms still relies on textual descriptions that are less interactive and do not provide personalized preference matching. This study aims to develop a web-based local perfume recommendation system that integrates the Content-Based Filtering method with Term Frequency-Inverse

Document Frequency (TF-IDF) weighting technique and Cosine Similarity algorithm to evaluate the similarity of user criteria against target products in the database. The research employed a software development approach consisting of problem identification, literature review, data collection, system design, implementation, and testing. The system was designed using flowcharts, use case diagrams, and entity relationship diagrams. The system processes user criteria input including aroma, occasion, gender, price, and concentration, extracts them into weight vectors, and then calculates the distance value (similarity) to produce product ranking classifications. The implementation results indicate that the application is capable of executing the main search functions, displaying catalogs, receiving preference inputs, and providing accurate automatic recommendation feedback. Precision testing on 20 search scenario samples produced 17 relevant data (True Positive) and 3 irrelevant data (False Positive), with an overall system accuracy of 85%. These findings indicate that the combination of TF-IDF and Cosine Similarity is feasible to be implemented as an interactive recommendation engine to help consumers choose perfumes in a more independent, measurable, and accurate manner, as well as to minimize the risk of purchasing errors (blind buys).

**Keywords:** local perfume; web; content-based filtering; TF-IDF; cosine similarity; recommendation system

## 1. PENDAHULUAN

Pada tahap awal, industri ini didominasi oleh segelintir merek global, namun kini ratusan produsen lokal bermunculan menawarkan racikan keharuman dengan tingkat kompleksitas komponen yang bervariasi. Pesatnya peluncuran produk baru ini menuntut adanya sistem informasi digital yang mumpuni untuk menjembatani komunikasi antara produsen dan konsumen secara efisien. Meskipun demikian, proses pemasaran dan eksplorasi parfum di berbagai platform digital saat ini masih memicu sebuah paradoks baru, yakni fenomena *information overload*[1]. Konsumen dihadapkan pada kesulitan besar dalam memfiltrasi informasi tekstual yang berlebihan untuk menentukan produk mana yang paling beresonansi dengan identitas personal serta preferensi psikologis mereka[2].

Akar dari permasalahan ini terletak pada sifat pemilihan parfum yang sangat subjektif dan mengandalkan pengalaman indrawi yang abstrak, sementara interaksi di platform digital menghilangkan kemampuan penciuman secara fisik. Konsumen dipaksa untuk bergantung sepenuhnya pada susunan deskripsi tekstual yang sering kali menggunakan terminologi puitis namun ambigu, sehingga menyulitkan mereka dalam membedakan detail teknis komposisi aroma[3]. Kondisi ini menyebabkan pengguna kesulitan menerjemahkan teks menjadi ekspektasi penciuman nyata, sehingga proses pemilihan produk beralih menjadi aktivitas coba-coba (*trial and error*)[4]. Lebih jauh lagi, ketidaksesuaian interpretasi ini melahirkan tingginya angka *blind buy*, yaitu tindakan membeli parfum tanpa mengetahui aromanya terlebih dahulu. Risiko ini tidak hanya berpotensi menimbulkan ketidakpuasan fungsional, tetapi juga memicu kerugian ekonomi langsung bagi konsumen apabila produk yang dibeli tidak sesuai ekspektasi[5].

Kehadiran teknologi sistem rekomendasi (*Recommendation System*) membuka peluang besar untuk mengatasi hambatan interaksi digital tersebut dengan menawarkan pendekatan pencarian yang lebih cerdas dan personal. Dalam bidang temu kembali informasi (*information retrieval*), teknik *Content-Based Filtering* (CBF) telah diakui keandalannya untuk merekomendasikan entitas berdasarkan kemiripan atribut kontennya dengan profil kriteria yang dicari oleh pengguna[6]. Pendekatan ini sangat relevan diaplikasikan pada industri wewangian karena setiap produk memiliki parameter spesifik yang dapat dikonversi menjadi representasi vektor matematis. Pada saat yang sama,

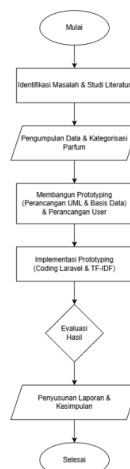
pemanfaatan algoritma pembobotan kata *Term Frequency-Inverse Document Frequency* (TF-IDF) memungkinkan sistem untuk secara otomatis menyeleksi dan memberikan nilai bobot tertinggi pada komponen aroma yang paling representatif di dalam suatu deskripsi produk[7]. Selanjutnya, algoritma *Cosine Similarity* digunakan karena kemampuannya menghitung jarak sudut kosinus antar-vektor, memberikan nilai kecocokan presisi tanpa terpengaruh oleh panjang deskripsi[8].

Penelitian-penelitian sebelumnya mengenai sistem rekomendasi telah banyak mengeksplorasi penggunaan algoritma *Cosine Similarity* pada berbagai platform digital [9], [10]. Namun, pendekatan konvensional tersebut umumnya hanya mengandalkan kemiripan teks secara *flat* tanpa mempertimbangkan prioritas kriteria pencarian dari sudut pandang psikologis konsumen [11]. Pada penelitian ini, diusulkan sebuah keunggulan baru (*novelty*) berupa implementasi kasta pembobotan hierarkis berlapis menggunakan metode TF-IDF. Sistem secara spesifik memberikan skor prioritas tertinggi pada elemen Aroma (5), disusul Gender (4), Occasion (3), Konsentrasi (2), dan Harga (1). Integrasi ini menutupi *gap* dari penelitian terdahulu karena sistem tidak hanya menghitung frekuensi kata, tetapi juga meniru alur kognitif konsumen yang memprioritaskan karakter wewangian di atas atribut komersial lainnya.

Penelitian ini bertujuan merancang dan mengimplementasikan aplikasi sistem rekomendasi parfum lokal berbasis web yang secara khusus memadukan *Content-Based Filtering*, pembobotan TF-IDF, dan *Cosine Similarity*. Aplikasi dirancang untuk mereduksi kompleksitas pencarian dengan mengekstraksi 7 variabel kriteria input pengguna menjadi urutan rekomendasi yang akurat. Secara praktis, aplikasi diharapkan mampu membantu konsumen melakukan penyaringan produk secara mandiri, terukur, dan terhindar dari kerugian *blind buy*, sekaligus memperkaya literatur rekayasa sistem informasi di sektor industri gaya hidup lokal[12].

## 2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan pengembangan perangkat lunak (*software development*) dengan kerangka kerja *Prototyping* untuk menghasilkan sistem rekomendasi parfum berbasis web. Sistem dirancang untuk menganalisis dan mencocokkan input kriteria dari pengguna terhadap dataset parfum lokal yang memiliki 7 parameter klasifikasi produk (Nama Parfum, Harga, Ukuran, Gender, Konsentrasi, Jenis Aroma, dan Occasion). Tahapan penelitian dirancang secara sistematis agar dapat dievaluasi secara iteratif pada setiap fase pengembangannya.



Gambar 1. Prosedur Penelitian

Pada Gambar 1, Flowchart prosedur penelitian tersebut menggambarkan tahapan kerja dalam pengembangan perangkat lunak yang dilaksanakan secara sistematis. Proses diawali dengan mengidentifikasi masalah dan melakukan studi literatur, dilanjutkan dengan pengumpulan serta kategorisasi data parfum. Tahap berikutnya adalah perancangan arsitektur sistem (prototyping) yang melibatkan pembuatan UML dan basis data. Setelah itu, sistem diimplementasikan ke dalam bahasa pemrograman menggunakan *framework* Laravel yang dilengkapi dengan kalkulasi algoritma TF-IDF dan *Cosine Similarity*. Langkah terakhir melibatkan evaluasi hasil menggunakan matriks pengujian sebelum akhirnya didokumentasikan dalam penyusunan laporan. Langkah – langkah yang dilakukan adalah:

### Identifikasi Masalah

Penelitian ini dimulai dengan mengidentifikasi permasalahan yang dialami konsumen saat mencari parfum di platform digital. Berdasarkan observasi, fenomena *information overload* dan keterbatasan pengalaman indrawi saat berbelanja *online* menyebabkan konsumen kesulitan menerjemahkan deskripsi tekstual wewangian, yang bermuara pada tingginya angka *blind buy*. Oleh karena itu, diidentifikasi kebutuhan mendesak akan sebuah mesin rekomendasi yang dapat memetakan preferensi secara objektif.

### Studi Literatur

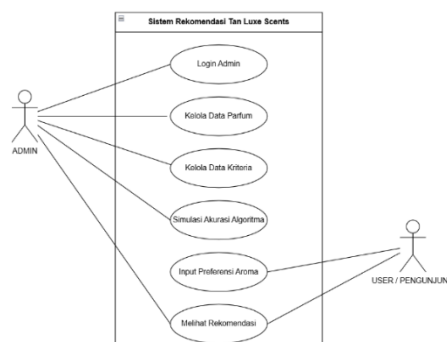
Peneliti melakukan kajian komprehensif dengan mencari sumber referensi dari berbagai artikel ilmiah dan publikasi akademik yang berkaitan dengan rekayasa sistem informasi [10]. Fokus kajian literatur ditujukan pada landasan teori *Content-Based Filtering*, teknik pengolahan teks berbahasa Indonesia, penerapan metode *Term Frequency-Inverse Document Frequency* (TF-IDF), dan logika matematika di balik algoritma pengukuran jarak vektor *Cosine Similarity*.

### Pengumpulan Data

Proses pengumpulan data dilakukan melalui ekstraksi spesifikasi produk secara internal dari *brand* parfum lokal Tan Luxe Scents. *Dataset* yang digunakan dalam pengujian sistem ini berjumlah 15 baris data parfum. Distribusi kategori data tersebut mencakup variasi aroma utama (seperti *floral*, *woody*, *citrus*), rentang *gender* (pria, wanita, *unisex*), hingga rekomendasi *occasion* (seperti formal atau harian) untuk memastikan variabilitas pengujian berjalan sesuai skenario perangkat lunak. Seluruh data diekstraksi ke dalam 7 variabel utama: nama produk, harga, ukuran, *gender*, konsentrasi, piramida aroma, dan rekomendasi situasi penggunaan.

### Perancangan sistem

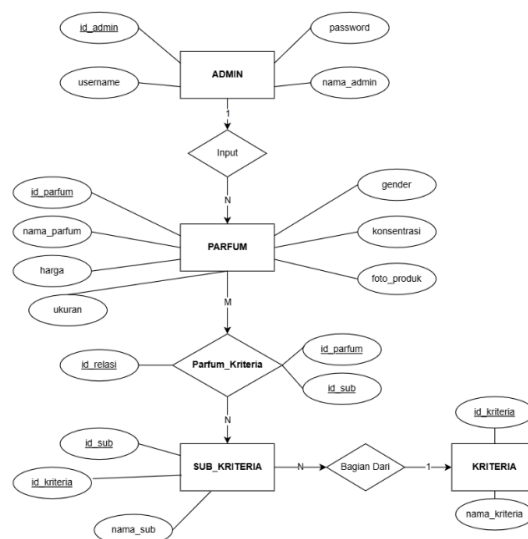
Perancangan arsitektur perangkat lunak dilakukan dengan menggunakan perangkat pemodelan *Unified Modeling Language* (UML). Pemodelan struktural ini bertujuan untuk memberikan cetak biru spesifikasi sistem yang mendefinisikan batas fungsionalitas dan interaksi antar-objek sebelum kode program ditulis[12].



Gambar 2. Use Case Diagram Sistem

Pada gambar 2 menjelaskan Use Case Diagram yang menggambarkan interaksi fungsional antara aktor dengan sistem rekomendasi. Aktor "Admin" memiliki otoritas penuh untuk masuk (*login*), mengelola katalog data parfum, serta melakukan simulasi internal. Di sisi lain, aktor "User" diizinkan berinteraksi dengan sistem tanpa login untuk memasukkan preferensi aroma dan segera melihat hasil *ranking* rekomendasi yang diproses oleh algoritma.

Setelah mendefinisikan interaksi fungsional melalui *Use Case Diagram*, tahapan krusial selanjutnya adalah membangun arsitektur penyimpanan data yang tangguh di sisi *back-end*. Proses ini diawali dengan mengidentifikasi seluruh entitas utama yang terlibat, mulai dari kredensial administrator, katalog produk parfum, hingga rincian sub-kriteria spesifik pendukung algoritma. Mengingat kompleksitas atribut parfum lokal—di mana satu varian produk dapat beririsan dengan banyak kriteria sekaligus (sebagai contoh, satu parfum memiliki kombinasi aroma *woody* dan *citrus*, serta direkomendasikan untuk *occasion* formal maupun santai)—maka diperlukan relasi data yang dinamis. Oleh karena itu, dirancanglah sebuah struktur tabel perantara (*pivot table*) untuk menangani relasi *Many-to-Many* agar pemrosesan pencarian menjadi efisien dan terhindar dari redundansi data. Pemetaan struktur tabel, penentuan atribut kunci (*Primary Key* dan *Foreign Key*), beserta kardinalitas relasi antar-entitas tersebut secara komprehensif divisualisasikan melalui *Entity Relationship Diagram* (ERD) berikut:



**Gambar 3.** Entity Relationship Diagram

Gambar 3 ERD memvisualisasikan arsitektur penyimpanan data relasional MySQL. Diagram ini menyoroti penggunaan tabel perantara (*Tabel\_Parfum\_Kriteria*) yang mengimplementasikan relasi *Many-to-Many* antara tabel master parfum dengan tabel rincian sub-kriteria. Desain relasional ini krusial untuk memastikan bahwa satu entitas parfum dapat menyimpan kombinasi multikriteria (misal: aroma *woody* sekaligus *citrus* untuk acara formal) secara fleksibel tanpa redundansi data.

### Algoritma TF-IDF

Fase komputasi sistem diawali dengan mengekstraksi nilai dari teks masukan pengguna melalui tahap *Text Preprocessing* (*Case Folding, Filtering, Stemming*). Selanjutnya, algoritma TF-IDF (*Term Frequency-Inverse Document Frequency*) digunakan untuk memberikan skor numerik dan mengekstraksi keunikan kata kunci. Pada sistem ini, pembobotan dioptimalkan dengan menerapkan kasta hierarki prioritas berlapis: Aroma

(5), Gender (4), Occasion (3), Konsentrasi (2), dan Harga (1). Rumus umum pembobotan TF-IDF dapat dituliskan sebagai berikut:

$$W_{t,d} = tf \times IDF \quad (1)$$

Nilai  $W_{t,d}$  merupakan bobot akhir dari suatu kata (*term*)  $t$  di dalam dokumen  $d$  [5]. Variabel  $tf$  (*term frequency*) adalah jumlah frekuensi kemunculan kata  $t$  pada kueri preferensi masukan pengguna. Sedangkan  $IDF$  (*Inverse Document Frequency*) adalah nilai inversi yang dihitung logaritmanya untuk merepresentasikan signifikansi atau keunikan kata  $t$  dalam keseluruhan *dataset* produk [6].

### Algoritma Cosine Similarity

Setelah nilai bobot vektor terbentuk melalui perhitungan TF-IDF, tahap berikutnya melibatkan iterasi pada matriks bobot dari vektor preferensi pengguna (A) dan vektor target produk parfum (B). Pada setiap iterasi dilakukan perhitungan perkalian titik (*dot product*) antar bobot kriteria. Nilai tersebut kemudian dinormalisasi dengan membaginya terhadap hasil perkalian panjang (*magnitude*) dari kedua vektor. Hasil perhitungan ini menentukan persentase kedekatan (*similarity*); jika kedua vektor memiliki atribut yang identik maka nilai kemiripan akan mendekati 1 (100%), sedangkan jika sama sekali berbeda akan bernilai 0. Setelah seluruh perhitungan selesai, nilai akhir diambil sebagai hasil utama berupa jarak kosinus, yang selanjutnya dimanfaatkan untuk mengklasifikasikan dan menentukan peringkat rekomendasi produk. Rumus umum algoritma *Cosine Similarity* dapat dituliskan sebagai berikut[8]:

$$Similarity(A,B) = \cos(\theta) = (A \cdot B) / (|A| \times |B|) \quad (2)$$

Pada Persamaan 2, nilai  $Similarity(A,B)$  atau  $\cos(\theta)$  merepresentasikan persentase kedekatan jarak sudut kosinus (bernilai 0 hingga 1) antara dua vektor [3]. Variabel  $A$  adalah matriks vektor pembobotan TF-IDF dari kriteria preferensi pengguna, sedangkan variabel  $B$  adalah matriks vektor atribut produk dalam basis data. Simbol  $(A \cdot B)$  menunjukkan hasil perhitungan perkalian titik (*dot product*) antar bobot kriteria, yang kemudian dinormalisasi dengan membaginya terhadap hasil perkalian panjang (*magnitude*) dari vektor  $A$  (simbol  $|A|$ ) dan vektor  $B$  (simbol  $|B|$ ).

Algoritma *Cosine Similarity* yang dikombinasikan dengan teknik pembobotan TF-IDF terbukti sangat baik digunakan dalam penghitungan tingkat kemiripan (*similarity*) antara preferensi pengguna dengan deskripsi produk, khususnya pada penelitian ini. Terlihat dari nilai evaluasi *Precision* sebesar 85%, di mana nilai tersebut merepresentasikan tingkat relevansi yang tinggi antara rekomendasi sistem dengan ekspektasi aktual pengguna. Secara keseluruhan, penerapan metode *Content-Based Filtering* dalam aplikasi sistem rekomendasi parfum berbasis web ini terbukti mampu memberikan saran pencarian yang objektif dan terukur. Dengan tingkat akurasi mencapai 85%, algoritma ini dapat dijadikan metode yang efektif untuk memeriksa dan menyeleksi kesesuaian profil keharuman berdasarkan input kriteria pengguna. Hasil tersebut mendukung fungsi utama aplikasi sebagai media penunjang keputusan yang interaktif dan adaptif dalam membantu konsumen memilih parfum lokal dengan cara yang efisien dan berbasis teknologi modern[1]. Selain itu, dalam implementasi sistem temu kembali informasi (*information retrieval*), integrasi dengan teknik pemrosesan teks mentah (*Text Preprocessing*) menjadi aspek penting dalam meningkatkan kualitas hasil perhitungan. Pendekatan ini memungkinkan sistem tidak hanya mencocokkan kata kunci secara harfiah, tetapi juga menganalisis signifikansi atribut produk—seperti *aroma*, *gender*, dan *occasion*—terhadap target pencarian secara lebih terstruktur dan kontekstual[6].

Penelitian ini memberikan kontribusi dalam pengembangan sistem *e-commerce* dan gaya hidup dengan pendekatan pemberian rekomendasi yang lebih terpersonalisasi. Secara umum, algoritma TF-IDF dan *Cosine Similarity* banyak dimanfaatkan dalam

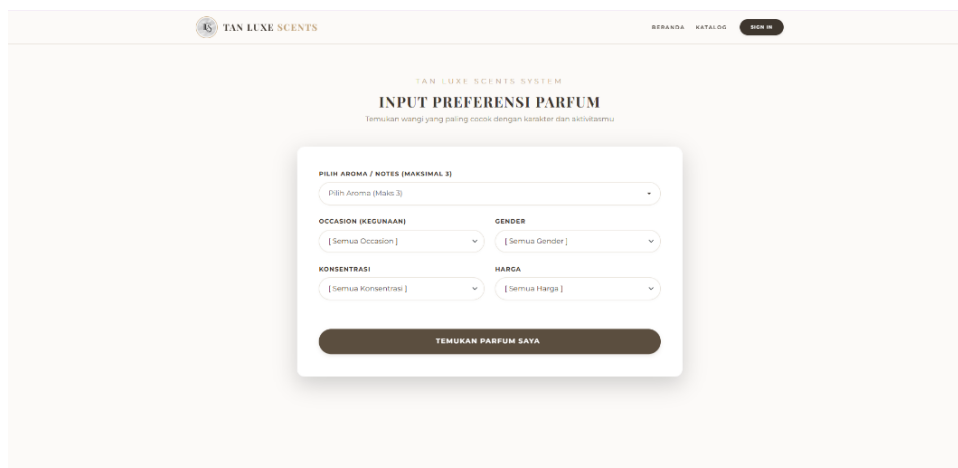
pengolahan teks, khususnya pada sistem pencarian dokumen akademik dan klusterisasi artikel untuk pemetaan topik[13], [14]. Namun, pemanfaatannya dalam konteks evaluasi kecocokan produk wewangian (parfum) masih relatif terbatas dan platform digital umumnya hanya menggunakan fitur penyaringan (*filter*) kategori yang statis. Oleh karena itu, penelitian ini mengusulkan penerapan algoritma *Cosine Similarity* yang dipadukan dengan kasta prioritas pembobotan TF-IDF untuk mengukur tingkat kemiripan antara kriteria pencarian dan spesifikasi parfum lokal[15]. Selain itu, penelitian ini juga mengembangkan aplikasi berbasis web yang mengintegrasikan elemen visual katalog produk dengan input preferensi pengguna secara langsung. Melalui pendekatan ini, sistem tidak hanya melakukan pencarian data dasar, tetapi juga memberikan evaluasi kecocokan berbasis nilai *similarity* dengan klasifikasi bertingkat, yaitu sangat direkomendasikan, direkomendasikan, dan kurang direkomendasikan. Dengan demikian, penelitian ini memperluas pemanfaatan algoritma TF-IDF dan *Cosine Similarity* dari domain pencarian teks umum menuju evaluasi kecocokan produk wewangian secara lebih objektif dan aplikatif guna menekan risiko *blind buy*.

### 3. HASIL DAN PEMBAHASAN

Hasil penelitian mencakup implementasi antarmuka sistem rekomendasi berbasis web, pengujian fungsionalitas fitur utama, serta simulasi perhitungan algoritma TF-IDF dan *Cosine Similarity* pada produk wewangian. Struktur aplikasi ini dirancang secara terintegrasi untuk mendukung alur pencarian produk: pengguna menginput preferensi melalui formulir dropdown, kemudian sistem memproses kriteria tersebut dan secara otomatis menampilkan daftar peringkat rekomendasi parfum berdasarkan nilai kemiripan tertinggi.

#### Implementasi Antarmuka Sistem

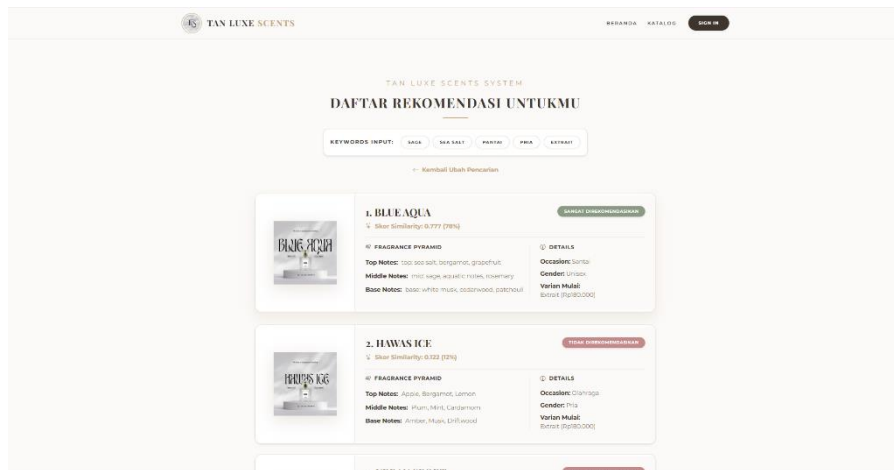
Desain antarmuka dirancang seringkasan mungkin menggunakan format formulir *dropdown* guna menghindari potensi salah ketik yang dapat mengacaukan pemrosesan algoritma. Pengguna dapat dengan mudah mendefinisikan parameter aroma, acara (*occasion*), jenis kelamin pengguna (*gender*), konsentrasi ketahanan, serta batasan harga.



**Gambar 4.** Halaman Form Preferensi

Berdasarkan Gambar 4, implementasi antarmuka *form* preferensi dirancang sebagai wadah interaktif untuk menangkap variabel kriteria dari konsumen. Halaman ini menggunakan elemen formulir terstandarisasi berupa *dropdown* pilihan ganda untuk parameter aroma, acara (*occasion*), jenis kelamin pengguna (*gender*), tingkat konsentrasi,

serta batasan harga. Pendekatan desain antarmuka ini bertujuan untuk memandu pengguna dalam memberikan masukan yang terstruktur, sekaligus meminimalisir kesalahan pengetikan (*typo*) yang berpotensi mengganggu presisi komputasi pada algoritma pencocokan.



**Gambar 5.** Halaman Hasil Rekomendasi

Selanjutnya pada Gambar 5, ditampilkan antarmuka hasil eksekusi dari algoritma pencarian berupa halaman daftar rekomendasi produk. Pada halaman ini, sistem menyajikan visualisasi data *output* dalam bentuk deretan kartu produk yang disusun secara menurun (*descending*) berdasarkan skor kemiripan algoritma. Produk yang memiliki tingkat kecocokan paling optimal diposisikan pada urutan paling atas (Peringkat 1), lengkap dengan lensa indikator visual "Sangat Direkomendasikan" beserta persentase skor *Cosine Similarity* untuk memudahkan pengguna dalam mengambil keputusan.

### Simulasi Komputasi Algoritma

Simulasi dilakukan secara manual untuk memverifikasi kebenaran logika kode. Misalkan seorang pengguna memasukkan kueri: "Floral, Kantor, Wanita, 30ml, Terjangkau, EDP". Sistem akan mencocokkannya dengan baris data parfum di pangkalan data, salah satunya adalah produk bernilai percontohan "Secret".

**Tabel 1.** Hasil Perhitungan Bobot TF-IDF

Term (t)	Q (tf)	Secret (tf)	df	IDF	Bobot Q (Ai)	Bobot Secret (Bi)
Floral	1	1	1	301	301	301
Kantor	1	1	1	301	301	301
Wanita	1	1	1	301	301	301
Secret	1	1	1	301	301	301
30ml	1	0	1	301	301	0
Terjangkau	1	1	1	301	301	301
EDP	1	1	1	301	301	301

Berdasarkan Tabel 1, proses perkalian *Term Frequency* (tf) dan inversnya mendemonstrasikan bagaimana mesin menyeleksi *term* yang tumpang tindih secara

absolut. Atribut 30ml bernilai nol pada kolom B<sub>i</sub> karena produk "Secret" tersebut tidak memiliki ukuran yang bersinggungan di dalam *database* komparasinya.

**Tabel 2.** Hasil Peringkat Rekomendasi (Cosine Similarity)

Peringkat	Nama Produk	Skor Cosine Similarity	Kategori Relevansi
1	Secret	0.926	Sangat Direkomendasikan
2	Orgs	0.815	Direkomendasikan
3	Scapium	0.756	Direkomendasikan
4	Blue Aqua	0.452	Kurang Direkomendasikan

Pada Tabel 2, eksekusi akhir *Cosine Similarity* menempatkan "Secret" pada kedudukan puncak (skor 0.926). Dominasi kecocokan pada kriteria bobot tinggi (Aroma Floral, Gender Wanita) membuat produk ini tetap menang dari varian lain, membuktikan formula matematis pembobotan berfungsi sesuai desain[9]. Karakteristik ini penting dalam konteks pengambilan keputusan, karena pengguna tidak sekadar memperoleh respons rekomendasi yang kaku, melainkan juga hasil penilaian kemiripan bertingkat yang menyajikan opsi alternatif produk secara lebih fleksibel.

### 3.3 Pengujian Sistem dan Analisis Presisi.

Sistem secara komprehensif diuji menggunakan teknik *Black Box Testing* untuk memastikan bahwa fitur-fitur pada antarmuka berjalan sempurna tanpa mempedulikan struktur kode internal. Pengujian memvalidasi bahwa sistem mampu mengeksekusi penyaringan logika harga dan menampilkan varian yang koheren tanpa *error*[13]. Tahap final melibatkan pengujian akurasi dengan metrik *Precision* menggunakan sampel evaluasi dari 20 responden berbeda.

**Tabel 3.** Rekapitulasi Hasil Pengujian Precision

Total Percobaan (Responden)	Sesuai / Relevan (True Positive)	Tidak Sesuai (False Positive)	Nilai Akurasi (Precision)
20	17	3	85%

Berdasarkan Tabel 3, hasil penelusuran menunjukkan bahwa 17 *output* kueri divalidasi sebagai luaran yang sangat presisi (*True Positive*), sementara 3 rekomendasi tidak relevan (*False Positive*) karena probabilitas penyimpangan prioritas absolut pengguna. Konversi metrik ini [ $TP / (TP + FP)$ ] membukukan angka akhir presisi sebesar 85%. Pencapaian ini merepresentasikan keandalan integrasi *Content-Based Filtering* yang sangat solid, menggeser perilaku coba-coba yang subjektif menjadi pengambil keputusan komputasional yang tangguh.

Kasus pada data kueri "Floral" dan "Kantor" yang dicocokkan dengan produk "Secret" pada Tabel 1 menghasilkan skor *similarity* sebesar 0.926 (92.6%). Hasil ini menunjukkan bahwa kesesuaian pada atribut dengan bobot kasta tinggi (Aroma dan *Occasion*) mampu mendongkrak nilai kedekatan vektor secara signifikan meskipun terdapat ketidakcocokan pada parameter minor seperti ukuran kemasan (30ml). Sebaliknya, jika terjadi ketidaksesuaian pada atribut aroma utama, nilai *similarity* akan turun drastis di bawah 0.500, yang menempatkan produk pada kategori kurang direkomendasikan. Dengan demikian, penerapan kasta pembobotan (Aroma 5, Gender 4, dst.) berpengaruh besar terhadap proporsi *similarity* yang dihasilkan system.

Dari sudut pandang implementasi algoritma, kombinasi TF-IDF dan *Cosine Similarity* terbukti sangat sesuai untuk data spesifikasi produk yang bersifat tekstual namun terstruktur seperti parfum. Setiap atribut dapat segera dipetakan ke dalam ruang vektor untuk dihitung jarak kosinusnya. Keunggulan ini memudahkan sistem menghasilkan rekomendasi yang terukur, stabil, dan cepat dieksekusi pada lingkungan *web* menggunakan *framework* Laravel. Hal tersebut sejalan dengan kebutuhan platform *e-commerce* yang membutuhkan respons instan agar pengalaman pengguna dalam mengeksplorasi produk tidak terganggu oleh kendala latensi server.

Hasil pengujian menunjukkan bahwa sistem mampu mengklasifikasikan tingkat kemiripan produk ke dalam kategori sangat direkomendasikan hingga kurang direkomendasikan secara konsisten. Temuan ini sejalan dengan penelitian sebelumnya yang memanfaatkan teknologi sistem rekomendasi dalam evaluasi pemilihan produk berbasis konten, di mana sistem mampu memberikan umpan balik otomatis dengan tingkat akurasi yang baik [7],[9]. Perbedaan utama pada penelitian ini terletak pada penggunaan kasta pembobotan hierarkis sebagai mekanisme filter ekonomi dan sosiologis (gender), sehingga penilaian tidak hanya bersifat pencocokan kata kunci, tetapi juga mempertimbangkan skala prioritas kebutuhan konsumen terhadap target produk parfum lokal.

Hasil pengujian juga memperlihatkan keterbatasan penting. Pertama, keakuratan rekomendasi sangat dipengaruhi oleh kelengkapan metadata produk di dalam basis data; data yang minim deskripsi akan menghasilkan skor yang kurang representatif. Kedua, *Cosine Similarity* bekerja pada representasi teks biner, bukan pada korelasi antar-aroma yang sebenarnya secara kimiawi. Akibatnya, dua aroma yang secara fisik berbeda namun memiliki label teks yang mirip dapat menghasilkan nilai kemiripan yang tinggi. Kasus ini mengindikasikan bahwa peningkatan di masa mendatang dapat diarahkan pada penggabungan pendekatan model pembelajaran mesin yang lebih sensitif terhadap profil kimiawi wewangian atau integrasi ulasan pengguna.

Jika dibandingkan dengan penelitian terdahulu, hasil penelitian ini memberikan kontribusi pada dua sisi sekaligus. Dari sisi pengembangan antarmuka, aplikasi menyediakan pengalaman pencarian yang lebih terarah dibanding metode filter konvensional di *marketplace*. Dari sisi komputasi, sistem tidak hanya berfungsi sebagai mesin pencari, tetapi juga menambahkan mekanisme pengukuran *similarity* untuk menghasilkan rekomendasi personal yang lebih bermakna. Dengan demikian, aplikasi tidak sekadar menyajikan katalog, melainkan juga berfungsi sebagai asisten belanja digital yang objektif.

Secara praktis, aplikasi dapat dimanfaatkan sebagai media pendamping bagi konsumen dalam menentukan pilihan sebelum melakukan transaksi nyata. Calon pembeli dapat melakukan simulasi pemilihan tanpa harus mencoba aroma secara fisik, yang secara langsung menekan angka kerugian akibat *blind buy*. Produsen parfum lokal juga memperoleh sarana untuk mempromosikan produk mereka secara lebih tertarget. Keunggulan utama penelitian ini terletak pada penyederhanaan proses pemilihan aroma yang kompleks menjadi alur komputasi yang ringan, mudah dipahami, dan dapat diakses secara luas melalui perangkat web.

Walaupun hasilnya menunjukkan kinerja yang baik, penelitian ini masih memiliki sejumlah keterbatasan. Dataset yang digunakan terbatas pada sejumlah produk parfum lokal tertentu dan belum mencakup seluruh variasi produk di pasar Indonesia. Selain itu, pengujian dilakukan pada jumlah responden yang terbatas sehingga generalisasi hasil perlu dilakukan secara hati-hati. Oleh sebab itu, penelitian lanjutan dapat diarahkan pada perluasan dataset, integrasi model *Collaborative Filtering* hibrida, pemakaian analisis

sentimen dari ulasan pelanggan, dan penambahan fitur perbandingan harga secara *real-time* agar nilai guna sistem bagi konsumen semakin meningkat.

## 4. KESIMPULAN

Penelitian ini telah berhasil membangun sebuah perangkat lunak sistem rekomendasi parfum lokal berbasis web dengan memadukan integrasi antarmuka intuitif dan evaluasi matematis secara otomatis. Perancangan sistem berbasis analisis kebutuhan yang direpresentasikan melalui *Use Case* dan diagram ERD terbukti mampu diwujudkan menjadi modul aplikasi dengan fungsionalitas yang solid menggunakan *framework* Laravel [16]. Pada aspek pemrosesan algoritmik, implementasi pembobotan hierarkis TF-IDF dan algoritma *Cosine Similarity* sukses menyaring dan mengkalkulasikan matriks atribut spesifik preferensi pengguna dengan tingkat akurasi yang tinggi. Validasi melalui pengujian pengguna menunjukkan performa yang impresif, di mana sistem menorehkan tingkat *Precision* mencapai 85%. Dengan demikian, pendekatan *Content-Based Filtering* ini dinyatakan sangat layak digunakan sebagai agen pemberi rekomendasi andal di sektor industri wewangian, secara proaktif membantu menjembatani keterbatasan indrawi konsumen serta memperkecil risiko kesalahan pembelian produk secara daring.

## 5. UCAPAN TERIMA KASIH

Penulis menyampaikan rasa terima kasih dan penghargaan yang mendalam kepada Program Studi Teknik Informatika, Fakultas Teknologi Informasi Universitas Hasyim Asy'ari. Secara khusus, apresiasi tertinggi ditujukan kepada para dosen pembimbing yang telah meluangkan waktu dan pikiran untuk memberikan arahan teknis yang tajam, serta kepada seluruh responden dan pihak terkait yang telah berpartisipasi menyumbangkan data krusial dalam mendukung kelancaran penyusunan riset perangkat lunak dan artikel ini.

## 6. REFERENCES

- [1] S. I. Adam, W. G. Mokodaser, F. I. Komputer, U. Klabat, and S. Utara, "IMPLEMENTASI SISTEM REKOMENDASI PRODUK E-COMMERCE MENGGUNAKAN CONTENT-BASED FILTERING BERBASIS COSINE," vol. 10, no. 2, pp. 418–424, 2025, doi: 10.51876/simtek.v10i2.1665.
- [2] Y. Imanita and M. D. Irawan, "Assessment Dalam Pemilihan Merek Parfum Berdasarkan Kepribadian," vol. VIII, pp. 159–168, 2025, doi: 10.53513/jursikb.v8i3.3754.
- [3] R. Priskila *et al.*, "Penerapan Metode Collaborative Filtering dan Content Based Filtering Pada Sistem Rekomendasi Smartphone Android," vol. 15, no. 3, pp. 477–487, 2024, doi: 10.31602/tji.v15i3.15255.
- [4] A. A. Firdausi, D. Hartanti, and A. A. Sari, "MENGGUNAKAN METODE CONTENT BASED FILTERING PADA PRODUK," vol. 10, no. 2, pp. 232–237, 2025, doi: 10.51876/simtek.v10i2.1569.
- [5] D. Frediansa, A. Mahmudi, and Y. A. Pranoto, "PREFERENSI ANTAR PEMBELI DENGAN MENGGUNAKAN COLLABORATIVE FILTERING BERBASIS WEB," vol. 9, no. 6, pp. 10587–10595, 2025, doi: 10.36040/jaiti.v9i6.69346.
- [6] I. Menggunakan, C. Algoritma, M. Suttan, H. Razhevva, and B. O. Lubis, "BARU

- UNIVERSITAS BSI MARGONDA FAKULTAS TEKNIK DAN," vol. 9, no. 2, pp. 2753–2758, 2025, doi: 10.36040/jaiti.v9i2.12849.
- [7] V. M. Hersianty *et al.*, "PENERAPAN ALGORITMA TF-IDF DAN COSINE SIMILARITY," vol. 9, no. 1, pp. 1619–1625, 2025, doi: 10.36040/jaiti.v9i1.51785.
- [8] D. Velamentosa and E. Zuliarso, "SISTEM REKOMENDASI FILM MENGGUNAKAN METODE CONTENT-BASED FILTERING," vol. 9, no. 2, pp. 2918–2922, 2025, doi: 10.36040/jaiti.v9i2.56716.
- [9] J. Wong, U. Multi, D. Palembang, K. Palembang, and S. Selatan, "IMPLEMENTASI TF-IDF DAN COSINE SIMILARITY UNTUK REKOMENDASI," vol. 10, no. 1, pp. 577–584, 2026, doi: 10.36040/jaiti.v10i1.71858.
- [10] I. Uml, S. Informasi, and P. Studi, "Impression : Jurnal Teknologi dan Informasi," vol. 4, no. 2, 2025, doi: 10.59086/jti.v4i2.902.
- [11] L. D. Nurmuthia, R. Imtiyaz, N. Tri, and M. Sagala, "ScienceDirect Personalized Perfume Recommendations Based on User Personalized Perfume Recommendations Based on User Descriptions Using Cosine and Jaccard Similarity Descriptions Using Cosine and Jaccard Similarity," *Procedia Comput. Sci.*, vol. 269, pp. 1097–1106, 2025, doi: 10.1016/j.procs.2025.09.051.
- [12] M. S. Ramadhan, "Desain Aplikasi Mobile Pesan Antar Makanan Dengan Fitur Inovatif Perhitungan Nutrisi Dan Rekomendasi Menu Berbasis UML," vol. 4, pp. 11731–11743, 2024, doi: 10.31004/innovative.v4i3.
- [13] A. S. Aqmarina, F. P. Aditiawan, and H. E. Wahanani, "SURABAYA MENGGUNAKAN METODE BLACK BOX TESTING DENGAN," vol. 8, no. 1, pp. 855–860, 2024, doi: 10.36040/jaiti.v8i1.35303.
- [14] F. Fadilillah and F. Amir, "Public Satisfaction Survey Information System Using Laravel Framework at Local Government Public Services Sistem Informasi Survei Kepuasan Masyarakat dengan Framework Laravel pada Pelayanan Publik Pemerintah Daerah," vol. 5, no. July, pp. 1074–1082, 2025, doi: 10.57152/malcom.v5i3.11372.
- [15] R. Hanun *et al.*, "IMPLEMENTASI CONTENT-BASED FILTERING," vol. 9, no. 2, pp. 3243–3250, 2025, doi: 10.29408/jit.v8i2.30540.
- [16] A. Azhar and L. Hadjaratie, "Implementasi Framework Laravel Dalam Pembuatan Sistem Pencatatan Notula Berbasis Website," vol. 01, pp. 45–51, 2023, doi: 10.21456/vol13iss1pp45-51.