PENERAPAN METODE HYBRID ALGORITMA DECISION TREE DAN TF-IDF UNTUK SISTEM REKOMENDASI PARFUM BERBASIS WEB (AROMANIA)

Muhammad Rayhan Havandi[1], Muhammad Rafi Akbar Azhari[2], Achmad Farlan Baihaqqi[3], Ahmad Zidan Alfarid[4], Giatika Chrisnawati[5]

[1,2,3,4,5]*Universitas Bina Sarana Informatika*

Email : rhvnd007@gmail.com[1], 17230085@bsi.ac.id[2], farlanronaldo77@gmail.com[3], ahmadzidan6.jidan@gmail.com[4], giatika@gcw.bsi.ac.id[5]

ABSTRAK

Perkembangan industri parfum yang pesat di Indonesia menghadirkan beragam pilihan produk bagi konsumen. Namun, banyaknya variasi aroma dan istilah teknis dalam dunia parfum seringkali menyebabkan kebingungan bagi calon pembeli dalam menentukan pilihan yang sesuai dengan preferensi mereka, fenomena yang dikenal sebagai information overload. Penelitian ini bertujuan untuk membangun sebuah sistem rekomendasi parfum berbasis web yang cerdas dan akurat dengan menerapkan metode Hybrid Filtering. Metode ini menggabungkan algoritma Decision Tree (C4.5) sebagai tahap penyaringan awal (filtering) berdasarkan atribut kategorikal (gender, harga, aktivitas), dan algoritma TF-IDF (Term Frequency-Inverse Document Frequency) serta Cosine Similarity sebagai tahap perankingan (ranking) berdasarkan kemiripan deskripsi aroma. Data yang digunakan terdiri dari 100 data parfum yang dikumpulkan dari berbagai sumber e-commerce. Hasil pengujian model menggunakan metode confusion matrix dengan data uji sebanyak 20 data menunjukkan bahwa sistem rekomendasi hybrid ini mampu mencapai tingkat akurasi sebesar 90%. Hasil penelitian membuktikan bahwa penggabungan kedua algoritma tersebut efektif dalam memberikan rekomendasi yang tidak hanya sesuai secara kategori tetapi juga memiliki relevansi konten yang tinggi terhadap preferensi pengguna.

Kata Kunci : *Sistem Rekomendasi, Parfum, Hybrid Filtering, Decision Tree, C4.5, TF-IDF, Cosine Similarity, Aplikasi Web*

ABSTRACT

The rapid growth of the perfume industry in Indonesia has presented consumers with a wide variety of product choices. However, the vast array of scents and technical terminology within the perfume world often causes confusion for potential buyers when determining choices that match their preferences, a phenomenon known as information overload. This study aims to develop an intelligent and accurate web-based perfume recommendation system by applying the Hybrid Filtering method. This method combines the Decision Tree (C4.5) algorithm as an initial filtering stage based on categorical attributes (gender, price, activity), and the TF-IDF (Term Frequency-Inverse Document Frequency) algorithm along with Cosine Similarity as a ranking stage based on scent description similarity. The dataset used consists of 100 perfume records collected from various e-commerce sources. Model testing results using the confusion matrix method with 20 test data points indicate that this hybrid recommendation system achieved an accuracy rate of 90%. The research results demonstrate that combining these two algorithms is effective in providing recommendations that are not only categorically appropriate but also possess high content relevance to user preferences.

Keywords : *Recommendation System, Perfume, Hybrid Filtering, Decision Tree, C4.5, TF-IDF, Cosine Similarity, Web Application.*

PENDAHULUAN

Pada zaman sekarang, parfum merupakan salah satu produk yang harus dimiliki oleh semua orang. Bukan hanya sebagai penambah aroma tubuh atau penghilang aroma yang tidak sedap, namun parfum sudah menjadi identitas bagi seseorang yang dapat membuat orang lain dapat mengenalinya hanya dari parfumnya saja. Parfum juga menjadi ciri khas seseorang kala berada dilingkungan terdekat maupun masyarakat. dalam pembagian parfum dibagi menjadi 8 kelompok yaitu floral, fresh, fruity, woody, sweet, musky, spicy dan citrus (Agus Setiawan Fadhlan et al., 2024).

Semakin banyak parfum yang beredar menjadikan masyarakat bingung untuk memilih jenis parfum yang tepat. Untuk mengatasi banyaknya jenis informasi yang membuat banyak orang bingung seperti ini dibutuhkan teknologi yang dapat berperan sebagai asisten pribadi. Sistem Pendukung Keputusan (SPK) atau Sistem Rekomendasi hadir sebagai solusi untuk memfilter ratusan pilihan dan memandu para pengguna menuju keputusan yang lebih baik. Sistem ini dapat menyederhanakan proses pemilihan yang kompleks menjadi serangkaian pertanyaan yang lebih mudah dipahami. Maka diperlukan perancangan sistem rekomendasi jenis parfum yang dapat mempermudah orang orang dalam mencari tahu apa jenis parfum yang cocok dengan kepribadiannya (Zulaikah et al., 2024).

Penelitian sebelumnya banyak menggunakan metode klasifikasi seperti Decision Tree untuk memberikan rekomendasi. Metode ini sangat efektif dalam menyaring produk berdasarkan kriteria tegas (kategorikal) seperti gender, rentang harga, atau ketahanan. Namun, Decision Tree memiliki keterbatasan dalam menangani preferensi yang lebih nuansa dan deskriptif, seperti kemiripan karakteristik aroma yang spesifik. Di sisi lain, metode Content-Based Filtering dengan algoritma TF-IDF (Term Frequency-Inverse Document Frequency) sangat unggul dalam mengukur kemiripan teks deskripsi, namun kurang efisien jika harus memproses seluruh database tanpa penyaringan awal.

Untuk mengatasi kelemahan masing-masing metode tersebut, penelitian ini mengusulkan pendekatan Hybrid. Penulis akan menggabungkan algoritma Decision Tree (C4.5) sebagai tahap awal (filtering) untuk mengklasifikasikan kandidat parfum yang relevan berdasarkan atribut dasar, dan kemudian menggunakan algoritma TF-IDF serta Cosine Similarity sebagai tahap akhir (ranking) untuk mengurutkan rekomendasi berdasarkan tingkat kemiripan deskripsi aroma. Pendekatan gabungan ini diharapkan mampu memberikan rekomendasi yang tidak hanya akurat secara kategori, tetapi juga sangat relevan secara konten aroma.

Berdasarkan latar belakang tersebut, penulis mengangkat judul penelitian "Penerapan Metode Hybrid Menggunakan Algoritma Decision Tree dan TF-IDF untuk Sistem Rekomendasi Parfum Berbasis Web (Aromania)".

Berikut ini adalah rumusan masalah berdasarkan latar belakang yang telah di uraikan sebelumnya yaitu :

1. Bagaimana Preferensi aroma parfum dapat dianalisis untuk memahami pola pembelian konsumen?
2. Apa saja faktor yang dapat mempengaruhi pengguna dalam menentukan jenis parfum ditengah tengah maraknya merk parfum yang beredar dikalangan masyarakat?
3. Bagaimana merancang sistem rekomendasi parfum yang menggabungkan algoritma Decision Tree (C4.5) untuk klasifikasi awal dan TF-IDF untuk perankingan relevansi?

Tujuan dari penelitian ini adalah sebagai berikut ini :

1. Membantu mencari tahu seberapa akurat algoritma decision tree dan TF-IDF bekerja dalam merekomendasikan parfum kepada pengguna.
2. Membantu masyarakat untuk mendapatkan parfum yang sesuai dengan kepribadiannya.
3. Membantu memberikan gambaran terhadap peneliti selanjutnya tentang pembuatan aplikasi serupa

Manfaat dari penelitian ini adalah sebagai berikut ini :

1. Meningkatkan pemahaman tentang parfum kepada masyarakat luas
2. Memudahkan proses pemilihan parfum kepada seluruh pengguna
3. Membantu menghemat waktu dalam mencari rekomendasi parfum
4. Meningkatkan penjualan dan promosi dari tiap brand

Berdasarkan penelitian yang telah kelompok kami buat dalam kegiatan ini, hipotesis yang kami dapatkan dalam penelitian ini adalah sebagai berikut :

H0 : Penerapan metode Hybrid (Decision Tree dan TF-IDF) tidak memberikan peningkatan relevansi yang signifikan dibandingkan metode tunggal dalam rekomendasi parfum.

H1 : Penerapan metode *Hybrid* (*Decision Tree* dan TF-IDF) efektif dan mampu memberikan rekomendasi parfum yang lebih relevan dan akurat dengan menggabungkan kekuatan klasifikasi dan pencocokan konten.

Penelitian ini difokuskan hanya pada keakuratan dalam penggunaan algoritma decision tree dan TF-IDF dalam memberikan rekomendasi parfum berdasarkan dataset yang telah kami buat sendiri dengan sedemikian rupa dengan variabel seperti gender, kategori harga, aktivitas, proyeksi, kelompok wangi, ketahanan rata-rata jam yang membuat hasil dari temuan tidak dapat langsung digunakan pada dataset yang lain walaupun dataset tersebut memiliki kemiripan yang cukup besar.

METODE

Tahapan Penelitian

Untuk memastikan bahwa penelitian yang kami lakukan berjalan sesuai dengan yang diharapkan dan berjalan secara terstruktur, penulis menyusun tahapan penelitian yang sistematis. Alur penelitian ini diawali dengan identifikasi masalah hingga evaluasi akhir. Berikut adalah rincian dari tahapan tahapan penilitian sebagai berikut :

1. Identifikasi Masalah : Tahapan awal ini dilakukan berdasarkan identifikasi permasalahan yang dihadapi oleh konsumen parfum, yaitu kesulitan dalam memilih varian yang sesuai akibat banyaknya pilihan (Information Overloaded) dan kompleksitas istilah wewangian dari sebuah parfum.
2. Pengumpulan data : Penulis melakukan pengumpulan dari data parfum secara manual dari berbagai sumber terpercaya. Data yang diambil termasuk kedalam data data seperti nama parfum, ketahanan, harga, proyeksi wangi (soft atau strong), dan aktivitas (outdoor, indoor atau fleksibel) yang semua ini dilakukan untuk memberikan atribut kategorikal kepada decision tree dan atribut deskriptif teks kepada TF-IDF.
3. Pra-pemrosesan data : terjadi 2 kegiatan ditahap ini yaitu Preprocessing Kategorikal yang dilakukan untuk membersihkan dan mengelompokkan data untuk klasifikasi. Dan Preprocessing Teks yang mana untuk melakukan Case Folding, Tokenizing, Stopword Removal, dan Stemming pada deskripsi aroma.

4. Penerapan Algoritma : Penerapan algoritma C4.5 disini berguna untuk menyaring kandidat parfum berdasarkan kriteria pengguna. Dan penerapan TF-IDF dan Consine Similarity pada kandidat hasil saringan untuk mengurutkan rekomendasi berdasarkan kemiripan deskripsi.
5. Perancangan sistem : Membuat rancangan sistem dari aplikasi Aromania juga harus dilakukan untuk mengetahui bagaimana alur atau cara kerja dari sistem menggunakan diagram UML (Use case dan Activity Diagram) serta perancangan antarmuka pengguna (User Interface).
6. Hasil & pengujian : Pengujian dari hasil rancangan diimplementasikan pada kode program seperti HTML atau JavaScript dan dilakukan testing data yang berguna untuk memvalidasi kinerja sistem.

Analisa Data

Analisa data merupakan komponen utama yang dibutuhkan dalam penelitian yang telah penulis bangun bersama. Dataset yang telah dikumpulkan akan digunakan secara spesifik untuk kebutuhan sistem rekomendasi dari aplikasi "Aromania".

Sumber Data

Sumber data yang digunakan dalam membangun dataset ini adalah data sekunder yang berjumlah 102 parfum. Sampel data yang digunakan mencakup berbagai macam merk parfum dari dalam maupun luar negeri.

Atribut Data

Atribut data yang digunakan pada dataset ini terdiri dari 7 atribut yang terbagi menjadi 6 atribut input (Kriteria) dan 1 atribut target (Label kelas). Seluruh atribut yang digunakan dalam dataset bersifat kategorikal agar mempermudah pembentukan cabang pada pohon keputusan.

Tabel 1. Atribut Data

No.	Nama Atribut	Tipe Data	Peran	Keterangan Nilai
1	Gender	Nominal	Input (fitur)	Men, Women, Unisex
2	Kategori harga	Ordinal	Input (fitur)	Murah, Menengah, Mahal
3	Aktivitas	Nominal	Input (fitur)	Indoor, Outdoor, Fleksibel
4	Proyeksi	Nominal	Input (fitur)	Soft, Strong
5	Kelompok wangi	Nominal	Input (fitur)	Floral, Woody, Fresh, Sweet, fruity, Spicy, Musky, Citrus
6	Ketahanan	Ordinal	Input (fitur)	Ringan, Sedang, Tahan lama
7	Nama Parfum	Nominal	Target (label)	Nama nama parfum rekomendasi

Pra-pemrosesan Data

Pra- pemrosesan data atau Preprocessing data merupakan sebuah tahapan yang dilakukan sebelum data digunakan untuk melatih model yang mana hal ini mencakup beberapa tahapan dikarenakan kita menggunakan metode hybrid yaitu sebagai berikut :

A Tahap Pertama :

1. Data Cleaning : Tahap ini merupakan tahapan yang dilakukan untuk memastikan tidak ada nilai yang kosong atau duplikat dalam dataset.
2. Generalisasi Atribut "kelompok wangi" : Menyederhanakan puluhan deskripsi yang didapatkan berdasarkan data yang diperoleh dari internet menjadi 8 kategori utama seperti fresh, woody, floral dan lainnya agar meningkatkan akurasi pengelompokan.
3. Diskretisasi atribut "Ketahanan" : Mengubah data numerik (jam) menjadi data kategori yang dapat digunakan dengan aturan

- Ringan : ≤ 4 jam

- Sedang : > 4 jam sampai dengan 8 jam
- Tahan lama : > 8 jam

B Tahap Kedua :

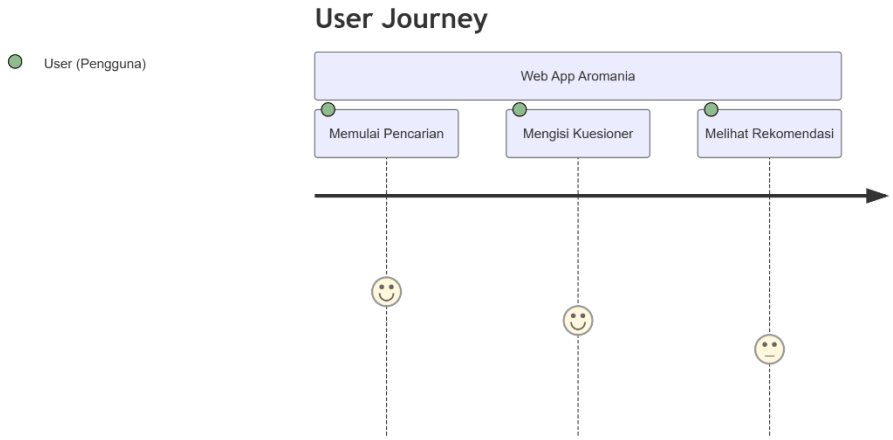
1. Case Folding: Mengubah seluruh huruf menjadi huruf kecil (lowercase) agar seragam (misal: "Vanilla" menjadi "vanilla").
2. Tokenizing: Memecah kalimat deskripsi menjadi potongan kata-kata tunggal (token).
3. Stopword Removal: Menghapus kata-kata umum yang tidak memiliki makna relevansi aroma, seperti "dan", "yang", "untuk", "dengan".
4. Stemming: Mengubah kata berimbuhan menjadi kata dasar (misal: "menyegarkan" menjadi "segar") untuk meningkatkan akurasi pencocokan kata kunci.

Rancangan Model (UML)

Perancangan sistem digambarkan menggunakan Unified Modeling Language (UML) untuk memvisualisasikan interaksi antara pengguna dan sistem. Berikut ini adalah perancangan dari Model aplikasi Aromania.

Use Case Diagram

Use Case Diagram Use Case diagram menggambarkan fungsionalitas sistem dari sudut pandang pengguna. Dalam sistem "Aromania", hanya terdapat satu aktor yaitu User (Pengguna).



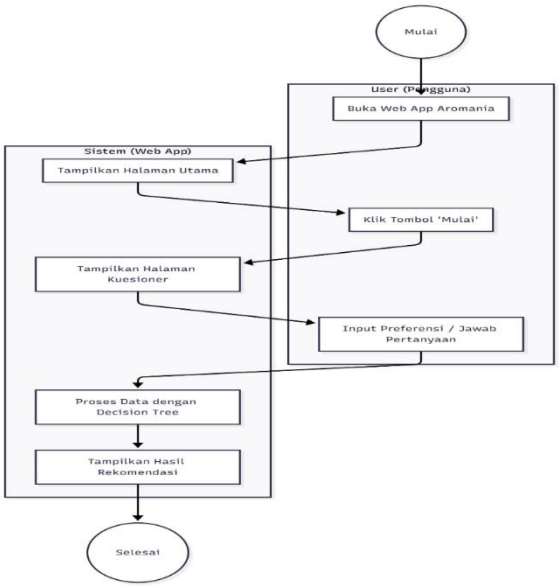
Gambar 1. Case Diagram

Keterangan :

1. Aktor : User (Pengguna Umum)
2. Use case 1 (Memulai Pencarian) : Pengguna mengakses sistem dan memulai pencarian parfum.
3. Use case 2 (Mengisi Kuisioner) : Pengguna menjawab beberapa pertanyaan yang telah disediakan dalam sistem.
4. Use case 3 (Melihat Rekomendasi) : Sistem menampilkan hasil rekomendasi parfum sesuai dengan preferensi pengguna.

Activity Diagram

Activity diagram menjelaskan alur kerja atau *Workflow* dari mulai pengguna membuka web applikasi Aromania hingga mendapatkan hasil rekomendasi sesuai dengan yang diinginkan.



Gambar 2. Activity Diagram

Penjelasan alur kerja :

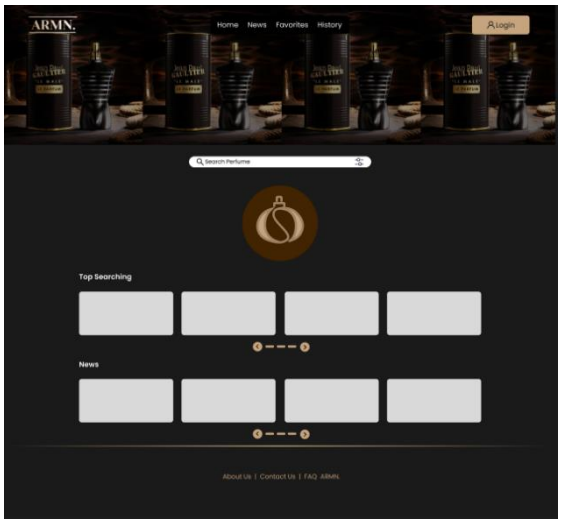
1. Pengguna membuka web ”Aromania”
2. Sistem menampilkan halaman utama (Landing page”
3. Pengguna menekan tombol “Cari parfum”
4. Sistem menampilkan formulir pertanyaan atau kuisioner berdasarkan atribut (Gender, wangi, ketahanan dan lain lainnya).
5. Pengguna memilih opsi jawaban sesuai dengan preferensi mereka
6. Sistem memproses input pengguna menggunakan aturan dari algoritma C4.5 untuk melakukan filter database dan Sistem menggunakan filtering dengan TF-IDF untuk memberikan rekomendasi yang terbaik
7. Sistem menampilkan hasil rekomendasi
8. Sistem selesai

Rancangan Antarmuka

Rancangan antarmuka atau biasa disebut *User Interface* merupakan sebuah tahapan yang menunjukkan bagaimana desain dari sebuah projek yang akan kita suguhkan kepada para pengguna. Rancangan antarmuka yang kami lakukan menggunakan tools Figma yang dapat memberikan gambaran visual aplikasi sebelum tahapan implementasi kode. Desain yang dibuat menggunkan konsep minimalis dan responsif.

Halaman Utama

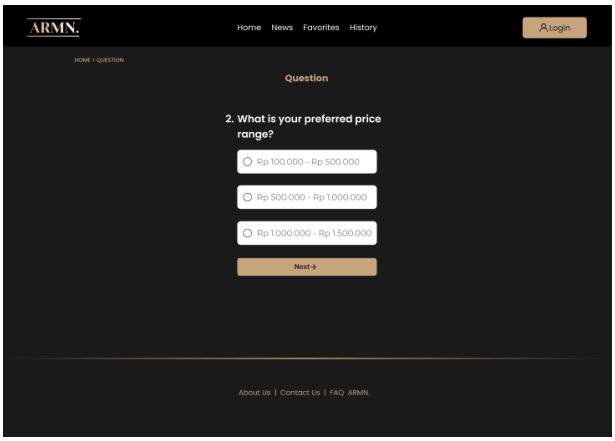
Halaman ini adalah tampilan awal saat para pengguna memasuki atau mengunjungi web yang telah kami buat. Halaman ini berisi nama aplikasi, deskripsi singkat, tombol search dan tombol memulai aplikasi.



Gambar 3. Halaman Utama

Kuisioner

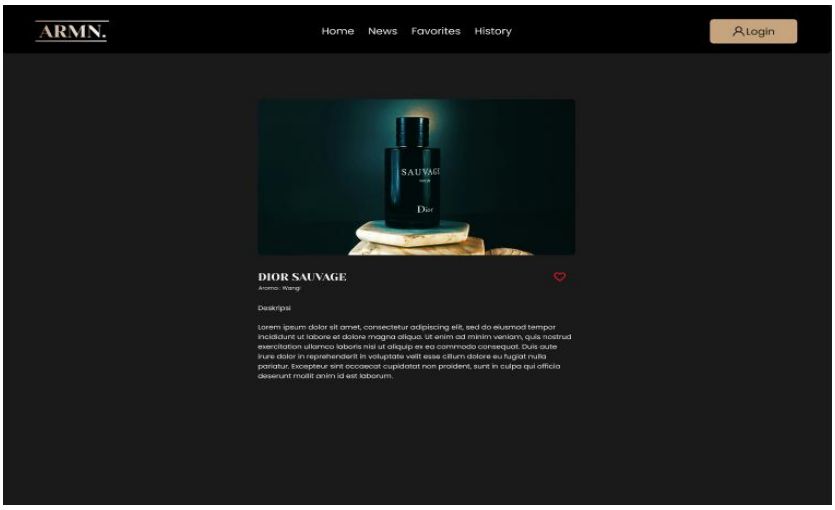
Halaman kuisioner dibuat sebagai page yang akan pengguna lihat agar bisa memberikan preferensi yang sesuai dengan mereka berdasarkan pertanyaan pertanyaan yang telah kami buat. Halaman ini hanya berisi pertanyaan yang telah kami buat untuk pengguna.



Gambar 4. Kuisioner

Halaman Hasil Rekomendasi

Rancangan Halaman Hasil Rekomendasi Halaman ini muncul setelah sistem selesai memproses input. Halaman ini akan menampilkan nama parfum rekomendasi, gambar visual produk, deskripsi singkat dan harga dan tombol mengulang pencarian.



Gambar 6. Rekaman Hasil Rekomendasi

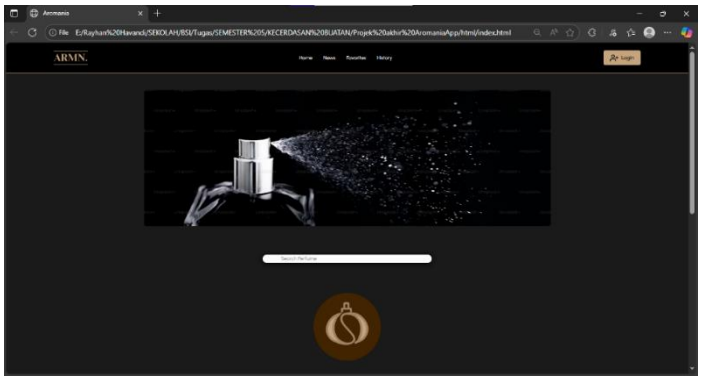
HASIL PENELITIAN DAN PEMBAHASAN

Implementasi

Pada tahapan sebelumnya, kita telah dijelaskan bahwa sistem rekomendasi dari web "Aromania" telah berhasil dibuat dengan baik dan mampu di implementasikan dalam bentuk aplikasi web. Implementasi ini juga mencakup user interface dan logika sistem yang menggunakan metode hybrid dari algoritma C4.5 dan juga algoritma TF-IDF.

Tampilan Halaman Utama

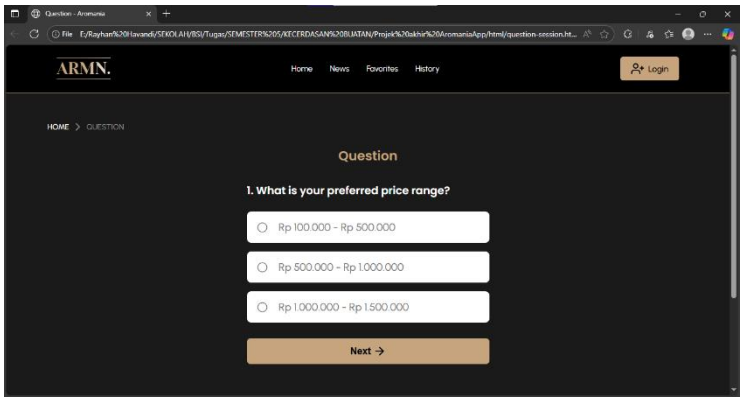
Halaman ini adalah tampilan awal saat para pengguna memasuki atau mengunjungi web yang telah kami buat. Halaman ini berisi nama aplikasi, deskripsi singkat, tombol search dan tombol memulai aplikasi.



Gambar 6. Tampilan Halaman Utama

Tampilan Halaman Kuisioner

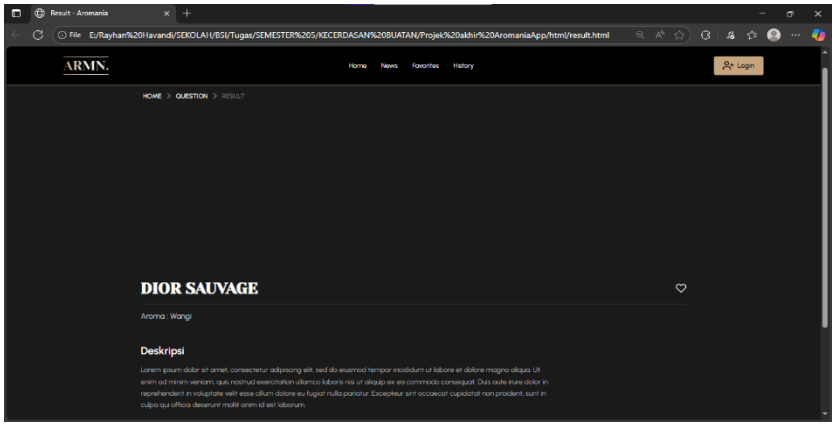
Halaman kuisioner dibuat sebagai page yang akan pengguna lihat agar bisa memberikan preferensi yang sesuai dengan mereka berdasarkan pertanyaan pertanyaan yang telah kami buat. Halaman ini hanya berisi pertanyaan yang telah kami buat untuk pengguna.



Gambar 7. Tampilan Halaman Kuisioner

Tampilan Halaman Hasil Rekomendasi

Rancangan Halaman Hasil Rekomendasi Halaman ini muncul setelah sistem selesai memproses input. Halaman ini akan menampilkan nama parfum rekomendasi, gambar visual produk, deskripsi singkat dan harga dan tombol mengulang pencarian.



Gambar 8. Tampilan Halaman Hasil Rekomendasi

Hasil Pengujian Model

Pada bagian ini, pengujian model bertujuan untuk melihat apa atau bagaimana model bekerja dengan tujuan untuk melihak kinerja dari model algoritma C4.5 dan juga TF-IDF dalam memberikan rekomendasi yang akurat. Untuk memvalidasi logika sistem hybrid, berikut disajikan simulasi perhitungan manual yang menggabungkan tahap klasifikasi dan perankingan.

Filtering dataset Dengan C4.5

Penelitian ini dilakukan dengan menggunakan dataset sebanyak 103 parfum yang dicari secara manual. Dataset dibagi menggunakan metode hold-out dengan rasio 80:20, yaitu :

1. Data latih : 80% dari data yang dimiliki digunakan untuk membangun model dari decision tree.
2. Data uji : 20% dari data yang dimiliki digunakan untuk menguji akurasi dari model.

Untuk memvalidasi logika algoritma yang diterapkan pada sistem, berikut disajikan perhitungan manual untuk satu iterasi (penentuan Root Node) menggunakan sampel 10 data latih. Berikut ini adalah contoh dari data latihnya :

Tabel 2. Penentuan Root Node

No	Gender	Harga	Aktivitas	Proyeksi	Wangi	Ketahanan	Nama parfum
1	Women	Menengah	Fleksibel	Soft	Sweet	sedang	VS Bombshell
2	Women	Mahal	Indoor	Soft	Sweet	Ringan	Black Opium
3	Women	Menengah	Indoor	Soft	Sweet	Sedang	Zara Nude
4	Men	Menengah	Fleksibel	Soft	Fruity	Sedang	Brodo Timur
5	Unisex	Menengah	Outdoor	Soft	Fresh	Sedang	Onix FWB
6	Women	Mahal	Fleksibel	Soft	Floral	Sedang	Chanel N5
7	Men	Mahal	Outdoor	Strong	Woody	Sedang	Bleu de Chanel
8	Women	Mahal	Outdoor	Soft	Floral	Lama	Dior J'adore
9	Unisex	Menengah	Indoor	Strong	Musky	Sedang	HMNS Darker
10	Unisex	Menengah	Fleksibel	Strong	Floral	Sedang	Saff&Co Irai

Langkah 1 : Menghitung Entropy Total

Diketahui jumlah kasus (S) = 10. Karena setiap data menghasilkan rekomendasi parfum yang unik (10 label berbeda), maka distribusi data sangat heterogen. Rumus Entropy:

$$Entropy(S) = \sum_{i=1}^{n} -p_i \log_2 p_i$$

Gambar 9. Rumus Entropy

Dengan perhitungan berdasarkan data sebagai berikut :

$$Entropy(Total) = 10 \times (-\frac{1}{10} \log_2 \frac{1}{10}) = 10 \times (0.1 \times 3.3219) = \mathbf{3.3219}$$

Gambar 10. Analisis Data

Langkah 2 : Menghitung Information gain untuk atribut “Gender”

Kita akan menghitung nilai Gain berdasarkan atribut dari “Gender” :

- Women (5 kasus) : 5 label unik : $-\log_2(1/5) = 2.3219$
- Men (2 kasus) : 2 label unik : $-\log_2(1/2) = 1.0000$
- Unisex (3 kasus) : 3 label unik : $-\log_2(1/3) = 1.5850$

$$Gain(S, Gender) = Entropy(S) - \sum \frac{|S_v|}{|S|} Entropy(S_v)$$

$$Gain(Gender) = 3.3219 - [(\frac{5}{10} \times 2.3219) + (\frac{2}{10} \times 1.0000) + (\frac{3}{10} \times 1.5850)]$$

$$Gain(Gender) = 3.3219 - [1.1609 + 0.2000 + 0.4755] = 3.3219 - 1.8364$$

$$Gain(Gender) = 1.4855$$

Gambar 11. Menghitung Atribut Gender

Berdasarkan hasil dari perhitungan yang telah kita lakukan, hasil dari Gain "Gender" adalah sebesar **1.4855**

Langkah 3 : Menghitung Information gain untuk atribut "Kelompok wangi"

Sama seperti sebelumnya, kita akan melakukan perhitungan Gain berdasarkan atribut "Kelompok wangi" :

- Sweet (3 kasus) : 3 label unik : $-\log_2(1/3) = 1.5850$
- Floral (3 kasus) : 3 label unik : $-\log_2(1/3) = 1.5850$
- Fruity, fresh, woody, musky : masing masing 1 label unik : karena hanya ada 1 data maka entropy menjadi 0

$$Gain(Wangi) = 3.3219 - [(\frac{3}{10} \times 1.585) + (\frac{3}{10} \times 1.585) + (\frac{1}{10} \times 0) \times 4]$$

$$Gain(Wangi) = 3.3219 - [0.4755 + 0.4755 + 0] = 3.3219 - 0.9510$$

$$Gain(Wangi) = 2.3709$$

Gambar 12. Analisis Kelompok Wangi

Berdasarkan hasil dari perhitungan yang telah kita lakukan, hasil dari Gain "Kelompok wangi" adalah sebesar **2.3709**

Jadi, berdasarkan perhitungan yang telah kita lakukan untuk mendapatkan nilai gain sebelumnya. Dapat kita ambil kesimpulan bahwa nilai gain dari kelompok wangi memiliki nilai yang lebih besar yaitu 2.3709 dibandingkan nilai gain dari gender dengan nilai 1.4855. Oleh karena itu, algoritma C4.5 memilih atribut Kelompok Wangi sebagai *Root Node* (Akar) karena memiliki kontribusi terbesar dalam mengklasifikasikan data.

Setelah proses training yang kita sudah lakukan. Sekarang kita masuk ketahapan hasil model dari decision tree yang berdasarkan dari 80 data latih. Berikut ini adalah representasi dari aturan atau *rules* yang dihasilkan :

Rule 1 : Wangi='Sweet' AND Harga='Menengah' AND Proyeksi='Soft' THEN Rekomendasi='VS Bombshell'

1. **Rule 2** : IF Wangi='Fresh' AND Gender='Men' AND Aktivitas='Outdoor' THEN Rekomendasi='Versace Dylan Blue'
2. **Rule 3** : IF Wangi='Woody' AND Ketahanan='Tahan Lama' THEN Rekomendasi='Aigner Black Man'
3. Hingga seterusnya untuk cabang cabang pada dataset yang kita miliki.

Tahap 2 : Ranking dengan TF-IDF & Cosine Similarity

Ketiga kandidat di atas kemudian diranking berdasarkan kemiripannya dengan *keyword* user: **"wangi vanilla manis"**.

Langkah 1: Text Preprocessing

- **Query (Q):** [wangi, vanilla, manis]
- **D1:** [aroma, vanilla, segar, sentuhan, bunga]
- **D2:** [wangi, buah, manis, ceria]
- **D3:** [dominasi, vanilla, manis, tahan, lama]

Langkah 2: Menghitung Bobot TF-IDF *(Asumsi IDF: wangi=1.4, vanilla=1.1, manis=1.1)*

Tabel 3. Menghitung Bobot TF-IDF

Kata	Query(Q)	D1(Parfum A)	D2(Parfum B)	D3(Parfum C)
wangi	1.4	0	1.4	0
vanilla	1.1	1.1	0	1.1
manis	1.1	0	1.1	1.1

Langkah 3 : Menghitung Cosine Similarity (Skor Kemiripan)

1. Skor D1 (Parfum A): Hanya cocok kata "vanilla". Skor $\approx$ 0.45
2. Skor D2 (Parfum B): Cocok kata "wangi" & "manis". Skor $\approx$ 0.78
3. Skor D3 (Parfum C): Cocok kata "vanilla" & "manis". Skor $\approx$ 0.85 (Paling Tinggi)

$$(Q \cdot D)/(|Q| \times |D|)$$

Kesimpulan: Sistem akan merekomendasikan Parfum C di urutan pertama karena mengandung dua kata kunci penting ("vanilla" dan "manis") yang sangat relevan dengan preferensi user, meskipun D2 juga memiliki skor tinggi. Nilai bobot pada simulasi ini menggunakan asumsi nilai IDF standar untuk penyederhanaan perhitungan manual.

Pengujian Hasil Akurasi

Setelah terbentuknya model yang telah kita buat sebelumnya, kita akan menggunakan 20 data uji untuk melihat Hasil prediksi model dibandingkan dengan label sebenarnya menggunakan *Confusion Matrix*.

Tabel 3. *Confusion Matrix*

	Prediksi benar	Prediksi salah
Hasil klarifikasi	18 (True positive)	2 (False negative)

Berdasarkan tabel diatas, sistem berhasil mendapatkan hasil 18 rekomendasi tepat dan mengalami hanya 2 kesalahan yang setelahnya akan kita hitung dengan rumus :

Pembahasan

Penerapan metode *Hybrid* terbukti efektif meningkatkan kualitas rekomendasi pada aplikasi web “Aromania”. Berikut ini adalah hasil dari pembahasan mengenai kualitas dari metode hybrid dalam sistem rekomendasi “Aromania” :

1. Efisiensi: Pada tahap ini, decision tree mampu memangkas ruang pencarian yang dapat membuat TF-IDF menjadi lebih optimal dalam mencari rekomendasi yang sesuai dengan pengguna inginkan.
2. Relevansi: Pada tahap TF-IDF, algoritma ini memberikan nuansa yang berbeda dengan menunjukkan hasil rekomendasi yang lebih personal kepada user. Dan hal itu ditunjukkan oleh sistem dengan mampu membedakan mana parfum manis segar dan mana parfum cerah manis.

$$Akurasi = \frac{\text{Jumlah Prediksi Benar}}{\text{Total Data Uji}} \times 100\%$$

$$Akurasi = \frac{18}{20} \times 100\% = \mathbf{90\%}$$

Gambar 13. Rumus Akurasi

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai penerapan metode *Hybrid* untuk sistem rekomendasi parfum "Aromania", maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Efektivitas Metode Hybrid: Penelitian ini berhasil memberikan sebuah nilai yang cukup tinggi dalam membutikkan bahwa metode hybrid dari algoritma C4.5 dan TF-IDF dapat bekerja dengan sinergis. Dengan C4.5 berperan sebagai filtering di awal dan TF-IDF yang berperan sebagai penentu peringkat yang memberikan rekomendasi spesifik dari preferensi yang diberikan oleh user
2. Implementasi Sistem: Sistem rekomendasi aromania berhasil memberikan hasil rekomendasi berdasarkan preferensi pengguna dengan memberikan pengalaman yang lebih fleksibel dan efektif dibandingkan dengan sistem yang lebih jadul dan manual
3. Akurasi Rekomendasi: Berdasarkan hasil pengujian yang dilakukan, sistem menunjukkan bahwa “Aromania” memberikan tingkat akurasi yang cukup tinggi yaitu 90%. Hal ini memberikan hipotesis alternatif atau H1 dapat diterima, yaitu model hybrid mampu memberikan hasil rekomendasi yang lebih akurat dengan meminimalisir rekomendasi yang tidak relevan sambil tetap mempertahankan personalisasi aroma

DAFTAR PUSTAKA

Agus Setiawan Fadhlan, Hamra, & Masnur. (2024). *Pola Pembelian Produk Parfum Menggunakan Algoritma Apriori Berdasarkan Data Mining Rule Asosiasi* (Vol. 4, Issue 3). https://jurnal.umpar.ac.id/index.php/sylog

Evicienna, & Amalia, H. (2013). ALGORITMA C4.5 UNTUK PREDIKSI HASIL PEMILIHAN LEGISLATIF DPRD DKI JAKARTA. In *Techno Nusa Mandiri* (Vol. 1). https://media.neliti.com/media/publications/227349-algoritma-c45-untuk-prediksi-hasil-pemil-ef30784a.pdf

Herdianto, E., Saepudin, S., & Sihabudin. (2022). SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN MERK PARFUM PADA WANITA MENGGUNAKAN METODE AHP.

Jurnal Sistem Informasi dan Teknologi Informasi), *4*(3), 99–105. https://jursistekni.nusaputra.ac.id/article/view/155/78

Juliartha, B. M. P., Fu'adi, A., & Ariani, D. F. Y. (2022). Analisa dan Rancangan Sistem Informasi Pariwisata Pacitan dengan UML dan ERD. *INFORMATION SYSTEM FOR EDUCATORS AND PROFESSIONALS*, *7*(1), 63–72. https://ejournal-binainsani.ac.id/index.php/ISBI/article/view/1920

Khairan Khairan, Nurul Husna, Hilda Maisyarah, & Muhammad Diah. (2025). *Formulation and Evaluation of Liquid Perfumes from Natural Fragrance Using Patchouli Oil*. *Vol. 12 Suppl. 2*, 119–129. https://jurnal.unpad.ac.id/ijpst/article/view/60116/25398

Rahmansyah, N., & Lusinia, S. A. (2021). *Buku Ajar SISTEM PENDUKUNG KEPUTUSAN: Vol. Vol 1* (Nugraha Rahmansyah & Armonitha Lusinia, Eds.; 1st ed.). Pustaka Galeri Mandiri. http://jurnal.pustakagalerimandiri.co.id

Sari, P., Syahputra, A., Zaky, N., Sibuea, R. U., & Zakhir, Z. (2022). *Perancangan Sistem Aplikasi Penjualan dan Layanan Jasa Laundry Sepatu Berbasis Website*. https://jurnal.ilmubersama.com/index.php/blendsains/article/view/67/26

Soebroto, A. A., Hasanah, U., Cholissodin, I., Sutrisno, & Febiola, Y. I. (2019). *Buku Ajar AI, Machine Learning & Deep Learning*. https://www.researchgate.net/publication/348003841

Zulaikah, D., Swanjaya, D., Kom, M., Rochana, S., Nusantara, U., & Kediri, P. (2024). Perancangan Sistem Rekomendasi Jenis Parfum dengan Metode Naive Bayes Classifier. In *Agustus* (Vol. 8). Online. https://proceeding.unpkediri.ac.id/index.php/inotek/article/view/4929