

---

# Large Language Model (LLM) untuk Rekomendasi Parfum

Gustian Herlambang<sup>\*1</sup>, Hanif Al Fatta<sup>2</sup>

<sup>1,2</sup>Magister Teknik Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta

Email: <sup>\*1</sup>[gustianher@students.amikom.ac.id](mailto:gustianher@students.amikom.ac.id), <sup>2</sup>[hanif.a@amikom.ac.id](mailto:hanif.a@amikom.ac.id)

(Naskah masuk: 6 Agustus 2025, diterima untuk diterbitkan: 20 Januari 2026)

**Abstrak:** Penelitian ini mengembangkan sistem rekomendasi parfum berbasis deskripsi teks menggunakan pendekatan Large Language Model (LLM). Model Sentence Transformer all-MiniLM-L6-v2 digunakan untuk merepresentasikan deskripsi dan notes parfum ke dalam bentuk vektor semantik, yang kemudian dibandingkan menggunakan cosine similarity. Sistem ini dievaluasi dengan membandingkan tiga pendekatan: TF-IDF, Word2Vec, dan LLM. Evaluasi dilakukan secara kuantitatif menggunakan skor cosine similarity dan precision, serta secara kualitatif melalui penilaian responden terhadap relevansi rekomendasi. Hasil menunjukkan bahwa meskipun Word2Vec mencatat skor cosine similarity rata-rata tertinggi (0,96) dibandingkan LLM dengan cosine similarity (0,68), LLM menghasilkan nilai precision tertinggi (0,67) dibandingkan Word2Vec (0,33) dan TF-IDF (0). LLM terbukti memberikan rekomendasi yang paling relevan secara semantik menurut persepsi pengguna. Keunggulan LLM terletak pada kemampuannya memahami konteks naratif dari deskripsi aroma, sehingga menghasilkan rekomendasi yang lebih personal. Penelitian ini menunjukkan bahwa LLM memiliki potensi tinggi dalam menangani tantangan sistem rekomendasi berbasis teks deskriptif, khususnya dalam domain parfum.

**Kata Kunci** – Sistem Rekomendasi;Parfum;Large Language Models;Sentence Transformer, Word Embedding

---

## Large Language Model (LLM) for Fragrances Recommendation

**Abstract:** This study developed a perfume recommendation system based on text descriptions using a Large Language Model (LLM) approach. The Sentence Transformer all-MiniLM-L6-v2 model was used to represent perfume descriptions and notes as semantic vectors, which were then compared using cosine similarity. The system was evaluated by comparing three approaches: TF-IDF, Word2Vec, and LLM. Evaluation was carried out quantitatively using cosine similarity and precision scores, as well as qualitatively through respondents' assessments of the relevance of the recommendations. The results show that although Word2Vec recorded the highest average cosine similarity score (0.96) compared to the LLM with a cosine similarity of 0.68, the LLM achieved the highest precision value (0.67) compared to Word2Vec (0.33) and TF-IDF (0). LLM provide the most semantically relevant recommendations according to user perception. The advantage of LLM lies in their ability to understand the narrative context of scent descriptions, resulting in more personalised recommendations. This study demonstrates that LLM hold great potential for addressing the challenges of text-based descriptive recommendation systems, particularly in the perfume domain.

**Keywords** – Recommender System;Fragrances;Large Language Model;Sentence Transformer;Word Embedding

---

### 1. PENDAHULUAN

Parfum merupakan barang yang saat ini sangat lekat dengan seseorang karena tidak hanya berfungsi sebagai penunjang penampilan, tetapi juga merepresentasikan karakter, suasana hati, dan preferensi personal seseorang [1]. Bagi sebagian orang, memilih parfum yang tepat menjadi bagian penting dari identitas diri [2]. Namun, di tengah maraknya produk parfum dengan variasi aroma, merek, dan karakteristik yang sangat beragam, seseorang kerap mengalami kebingungan dalam menentukan pilihan yang sesuai dengan selera atau kebutuhannya. Hal ini menjadi tantangan tersendiri, terutama bagi pengguna baru atau mereka yang tidak memiliki pengetahuan mendalam tentang komposisi aroma yang ada didalam produk parfum itu sendiri [3].

Dalam beberapa tahun terakhir, sistem rekomendasi telah berkembang pesat dan menjadi solusi teknologi yang banyak diterapkan dalam berbagai bidang untuk membantu pengguna menemukan produk atau layanan yang sesuai [4], [5], seperti dalam layanan agen, streaming, e-

commerce, hingga konten berita [6], [7]. Sistem rekomendasi bekerja dengan menganalisis preferensi pengguna, baik secara eksplisit maupun implisit, untuk menyarankan item yang relevan [8]. Dalam konteks pemilihan parfum, sistem rekomendasi berpotensi menjadi solusi yang sangat membantu dalam memberikan saran personalisasi berdasarkan preferensi aroma [9].

Meskipun demikian, membangun sistem rekomendasi untuk parfum membutuhkan step ekstra karena perlunya memahami konteks bahan yang diuraikan dalam bentuk teks. Tentunya berbeda dengan film atau produk elektronik yang memiliki atribut kuantitatif yang cenderung mudah diolah, parfum dideskripsikan dengan bahasa yang bersifat kualitatif dan subjektif, seperti “manis”, “hangat”, “mewah”, atau “elegan” [10]. Bahasa subjektif tersebut bahkan tercipta dari kombinasi pembentuk aroma atas, tengah, hingga dasar [11]. Penelitian sebelumnya menggunakan pendekatan metode berbasis keputusan multikriteria yang dilakukan oleh [2], [12], [13] namun tidak menangkap makna konteks dan kemiripan semantik teks. Sedangkan sistem rekomendasi berbasis representasi teks telah dilakukan oleh [14], [15] yaitu menggunakan metode *Latent Semantic Analysis* (LSA) dan *Word2Vec*. Kedua metode tersebut belum dapat menangkap konteks semantik dengan baik karena memiliki keterbatasan dimana hanya dapat menangkap makna semantik secara kontekstual dalam skala kalimat atau paragraf.

Seiring dengan perkembangan teknologi kecerdasan buatan, khususnya dalam bidang pemrosesan bahasa alami (*Natural Language Processing/NLP*), muncul pendekatan berbasis *Large Language Models* (LLM) yang mampu memahami dan mengolah bahasa secara kontekstual [16]. *Large Language Model* diperkenalkan pada tahun 2017 oleh [17], dimana konsep dasarnya adalah memperkenalkan *transformer* yang menggantikan arsitektur sebelumnya seperti RNN dan LSTM. Arsitektur LLM awal adalah BERT (*Bidirectional Encoder Representations from Transformers*) dan arsitektur GPT telah membuktikan efektivitasnya dalam berbagai tugas NLP.

Salah satu penerapannya adalah representasi teks dalam bentuk vektor (*embedding*) yang memungkinkan sistem memahami makna suatu kalimat, bukan hanya kata per kata [18]. Dalam konteks rekomendasi parfum, pendekatan ini membuka peluang untuk melakukan pencocokan deskripsi aroma secara semantik, sehingga rekomendasi yang dihasilkan menjadi lebih relevan dan personal. Penelitian terkait telah dilakukan oleh [3], [19] dimana menggunakan atribut deskripsi 855 hingga 20000 produk parfum. Namun hanya 62 hingga 100 produk parfum dengan 255 notes bahan yang berhasil dikurasi dan digunakan. Menggunakan *Large Language Models* peneliti menggunakan tiga *transformer* (MPNet-base, DistilRoBERTa dan MiniLMs). Didapatkan performa akhir sebesar 72 - 79%. Hal ini kurang memuaskan mengingat dataset yang digunakan hanya berkisar 100 produk.

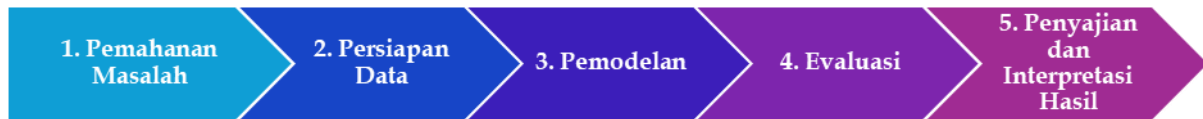
Penelitian ini bertujuan untuk membangun sebuah sistem rekomendasi parfum yang mampu memberikan saran berdasarkan deskripsi aroma menggunakan pendekatan *embedding* semantik berbasis LLM. Sistem ini memanfaatkan model bahasa *pre-trained* untuk menghasilkan representasi vektor dari deskripsi parfum, dan menggunakan kemiripan semantik untuk menentukan kedekatan antar deskripsi. Dengan demikian, ketika diberikan *input* berupa pertanyaan berdasarkan kondisi atau deskripsi keadaan atau preferensi aroma, sistem dapat merekomendasikan parfum yang memiliki karakteristik yang cocok secara kontekstual.

Kontribusi utama dari penelitian ini terletak pada penerapan pendekatan LLM dalam domain rekomendasi parfum, yang masih kurang dieksplorasi dalam studi sebelumnya. Tidak seperti pendekatan berbasis kata kunci, sistem ini mampu menangkap makna mendalam dari teks deskriptif aroma. Selain itu, penelitian ini juga menyajikan evaluasi sistem secara kuantitatif dan kualitatif untuk menunjukkan efektivitas pendekatan yang diusulkan dalam meningkatkan relevansi rekomendasi parfum berbasis teks.

## 2. METODE PENELITIAN

Penelitian ini bertujuan untuk membangun sistem rekomendasi parfum berbasis deskripsi aroma menggunakan pendekatan *embedding* semantik. Sistem ini dikembangkan dengan pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*) [20] yang mencakup enam tahap utama: pemahaman masalah, pemahaman data, persiapan data, pemodelan, evaluasi, dan

penyajian hasil. Alur proses CRSIP-DM ditunjukkan pada Gambar 1. Tahapan-tahapan tersebut diadaptasi secara sederhana dan diuraikan pada poin-poin di bawah ini disesuaikan dengan konteks penelitian ini.



Gambar 1. Alur Penelitian

### 2.1. Pemahaman Masalah

Pada tahap ini, ditetapkan bahwa permasalahan utama adalah bagaimana memberikan rekomendasi parfum yang relevan secara semantik berdasarkan deskripsi aroma. Pengguna sering kali kesulitan memilih parfum karena karakteristik aroma yang bersifat subjektif dan deskriptif. Oleh karena itu, pendekatan berbasis pemrosesan bahasa alami (*Natural Language Processing*) dipilih untuk menangani permasalahan ini.

### 2.2. Pemahaman dan Persiapan Data

*Dataset* yang digunakan dalam penelitian ini adalah *Perfume Recommendation Dataset* berasal dari repositori publik bernama Kaggle (<https://www.kaggle.com/datasets/nandini1999/perfume-recommendation-dataset>). *Dataset* terdiri dari 2191 entri parfum dan berbahasa Inggris. Setiap entri berisi kolom Nama Parfum, Merek, *Notes* atau Aroma, Deskripsi Aroma dan Url Gambar. Kolom yang digunakan untuk pelatihan model adalah kolom “Deskripsi” dan “Notes”. Deskripsi ini bersifat teks bebas yang mencerminkan karakteristik aroma, sedangkan kolom *Notes* berisi teks yang menjelaskan bahan baku aroma seperti jeruk bergamot (*bergamot*), kapulaga (*cardamom*), akar wangi (*vetiver*) dan sebagainya.

Sebelum digunakan dalam proses pemodelan, dataset dilakukan proses reduksi data sebanyak 80 baris dikarenakan terdapat data kosong didalamnya. Sehingga data yang digunakan untuk pelatihan sebanyak 2111 data. Selanjutnya, kolom deskripsi dan notes dipra-proses dengan langkah-langkah seperti mengubah teks menjadi huruf kecil (*lowercasing*), menghapus karakter non-alfabet, dan memastikan keseragaman struktur kalimat. Tidak dilakukan *stemming* maupun penghapusan *stopwords*, mengingat model bahasa yang digunakan sudah dilatih dengan struktur kalimat lengkap.

### 2.3. Pemodelan

Representasi semantik dari setiap deskripsi dan notes parfum dibentuk menggunakan model *Sentence Transformer*, yaitu all-MiniLM-L6-v2, yang merupakan varian ringan dari arsitektur BERT. Model ini menghasilkan vektor *embedding* berdimensi 384 untuk setiap deskripsi teks. Proses ini memungkinkan sistem memahami makna deskriptif dari aroma parfum secara lebih mendalam dan kontekstual.

Setelah mendapatkan vektor *embedding*, sistem menghitung kemiripan antar parfum menggunakan metode *cosine similarity*. Ketika pengguna memberikan input berupa deskripsi tertentu atau preferensi aroma, sistem akan menampilkan 3 parfum teratas yang memiliki nilai kemiripan tertinggi dengan input tersebut.

### 2.4. Evaluasi

Evaluasi dilakukan untuk menilai performa dari sistem rekomendasi parfum yang dikembangkan menggunakan tiga pendekatan berbeda, yaitu TF-IDF (*Term Frequency-Inverse Document Frequency*), Word2Vec, dan *Large Language Model* (LLM). Proses evaluasi ini mencakup

dua pendekatan utama, yakni evaluasi kuantitatif berbasis metrik dan evaluasi kualitatif berbasis persepsi pengguna.

Evaluasi Kuantitatif dilakukan dengan metode Metrik *cosine similarity*. Dimana metrik tersebut digunakan untuk mengukur tingkat kemiripan antara vektor representasi parfum input dengan vektor parfum-parfum yang direkomendasikan. Semakin tinggi nilai *cosine similarity*, semakin mirip secara semantik deskripsi antar parfum tersebut.

Evaluasi kedua yang dilakukan adalah Evaluasi Kualitatif, dimana bertujuan untuk menangkap persepsi subjektif dari pengguna terhadap hasil rekomendasi. Sebanyak tiga responden dengan latar belakang pengguna aktif parfum atau kolektor parfum. Hal tersebut penting mengingat kolektor atau pengguna parfum akan sangat familiar dengan notes parfum agar dapat menentukan kebenaran dari rekomendasi yang diberikan. Prosesnya adalah responden diminta untuk memberikan penilaian terhadap tingkat relevansi dan kesesuaian rekomendasi parfum yang dihasilkan oleh masing-masing metode. Penilaian dilakukan menggunakan skala 1 dan 0, di mana skor 1 berarti "relevan" dan skor 0 berarti "tidak relevan". Penilaian ini bertujuan untuk mengetahui sejauh mana sistem rekomendasi dapat menangkap preferensi atau nuansa aroma yang dirasakan pengguna. Evaluasi berbasis Metrik Presisi juga dilakukan untuk mengukur sejauh mana sistem rekomendasi cocok dengan preferensi pengguna atau responden.

### 2.5. Penyajian dan Interpretasi Hasil

Hasil evaluasi dan contoh output sistem disajikan dalam bentuk tabel. Analisis dilakukan untuk menjelaskan mengapa sistem merekomendasikan parfum tertentu berdasarkan kata kunci atau pola semantik dalam deskripsi.

## 3. HASIL DAN PEMBAHASAN

Dataset final yang digunakan pada penelitian sebanyak 2111 entri data dengan atribut "Deskripsi" dan "Notes" dalam bahasa Inggris. Proses prapemrosesan dilakukan untuk menghapus karakter khusus dan *stopwords*. Untuk mengukur relevansi rekomendasi, digunakan metrik *cosine similarity* terhadap deskripsi input dari pengguna. Top-3 parfum dengan nilai *similarity* tertinggi diambil sebagai hasil rekomendasi. Hasil evaluasi yang dilakukan terhadap tiga metode sistem rekomendasi parfum yaitu TF-IDF, Word2Vec, dan LLM (*Large Language Model*) menunjukkan hasil yang menarik terkait kinerja masing-masing model.

### 3.1. Analisis Performa

Performa ketiga metode dalam melakukan proses *training* data ditunjukkan pada Tabel 1. Data diukur berdasarkan waktu proses pelatihan data.

Tabel 1. Hasil Performa Pelatihan Model

Metode	Waktu Proses Pelatihan (s atau detik)
TF-IDF	0s
Word2Vec	10s
LLM	240s

Pada Tabel 1 ditunjukkan proses TF-IDF cenderung lebih cepat yaitu 0 detik, *Word2Vec* 10 detik dan yang paling membutuhkan waktu adalah LLM yaitu 4 menit. Proses TF-IDF lebih cepat karena hanya melibatkan perhitungan frekuensi kata dan tidak menggunakan pembelajaran mesin/*neural network*. Sedangkan *Word2Vec* mulai masuk ke ranah *machine learning* dengan

menggunakan *neural network* ringan untuk melatih vektor representasi dari kata-kata berdasarkan konteks. LLM adalah model *deep learning* kompleks yang sangat berat dan lambat dalam pelatihan karena melibatkan parameter yang jauh lebih banyak dan menggunakan arsitektur *transformer* yang jauh lebih kompleks.

### 3.2. Analisis Kuantitatif dengan Cosine Similarity

Berdasarkan pengujian awal (*baseline*) yang telah dilakukan dengan memasukkan inputan yaitu “*Saya ingin parfum yang tercium wangi bunga namun segar*” didapatkan hasil nilai *cosine similarity* rekomendasi yang ditunjukkan pada Tabel 2.

Tabel 2. Nilai *Cosine Similarity* Rekomendasi Berdasarkan Hasil Input Deskripsi

<i>Input</i>	<i>Metode/Model</i>	<b>Rekomendasi 1 (Nama Parfum   Nilai <i>Cosine Similarity</i>)</b>	<b>Rekomendasi 2 (Nama Parfum   Nilai <i>Cosine Similarity</i>)</b>	<b>Rekomendasi 3 (Nama Parfum Nilai )</b>
Saya ingin parfum yang tercium wangi bunga namun segar	TF-IDF	Orto Parisi Stercus Parfum   0.20	Serendipity 3 Serendipitous Eau de Parfum   0.12	Fueguia 1833 Cactus Azul Eau de Parfum   0.12
	Word2Vec	Kai - Eau de Parfum Eau de Parfum Spray   0.97	Kai - Perfume Oil Roll-on Perfume Oil   0.97	Monyette Paris Eau de Parfum Spray   0.96
	LLM	J-Scent Usubeni (Rosy Cheeks) Eau de Parfum   0.69	Xerjoff Casamorati Regio Eau de Parfum   0.68	Fueguia 1833 Cactus Azul Eau de Parfum   0.68

Nilai *cosine similarity* untuk model dengan TF-IDF berkisar antara 0.12 hingga 0.20, model Word2Vec memiliki skor *cosine similarity* paling tinggi berkisar antara 0.96 hingga 0.97, cukup berbeda dengan Large Language Model (LLM) yang hanya memiliki 0.68 hingga 0.69 saja. Namun nilai *cosine similarity* tidak menjamin hasil rekomendasinya sesuai dengan preferensi. Pada Tabel 3, dapat dilihat hasil rekomendasi parfum beserta dengan aroma atau notes yang ada.

Tabel 2. Rekomendasi Berdasarkan Hasil Input Deskripsi

<i>Input</i>	<i>Metode/Model</i>	<b>Rekomendasi 1 (Nama Parfum   Notes)</b>	<b>Rekomendasi 2 (Nama Parfum   Notes)</b>	<b>Rekomendasi 3 (Nama Parfum   Notes)</b>
Saya ingin parfum yang tercium wangi bunga namun segar	TF-IDF	Orto Parisi Stercus Parfum   Tembakau; <i>castoreum</i> ; kayu; kulit	Serendipity 3 Serendipitous Eau de Parfum   Coklat;Vanila Tahiti;Jeruk Purut	Fueguia 1833 Cactus Azul Eau de Parfum   bunga kaktus;kayu cedar; daun <i>mint</i>
	Word2Vec	Kai - Eau de Parfum Eau de Parfum Spray   bunga kacapiring ( <i>gardenia</i> ); bunga	Kai - Perfume Oil Roll-on Perfume Oil   bunga kacapiring ( <i>gardenia</i> ); bunga putih (melati, mawar)	Monyette Paris Eau de Parfum Spray   bunga kacapiring ( <i>gardenia</i> ); bunga muguet peranis; jeruk;

		putih (melati, mawar)		vanila;bunga putih (melati, mawar)
	LLM	J-Scent Usubeni (Rosy Cheeks) Eau de Parfum   jeruk mandarin;daun hijau;buah aprikot;bunga kenanga	Xerjoff Casamorati Regio Eau de Parfum   bunga lemon;jeruk kalabrian;bunga lavender	Fueguia 1833 Cactus Azul Eau de Parfum   bunga kaktus;kayu cedar; daun <i>mint</i>

Berdasarkan Tabel 3, TF-IDF memberikan rekomendasi yang kurang tepat. Input meminta parfum dengan “wangi bunga nan segar” namun justru diberikan rekomendasi parfum beraroma tembakau, kayu, vanila, bahkan coklat. Sedangkan Word2Vec dengan nilai *cosine similarity* tertinggi, justru dominan merekomendasikan parfum dengan aroma nuansa “bunga” saja. Disini *Large Language Model* (LLM) memberikan pemahaman semantik lebih baik karena berhasil merekomendasikan sesuai perintah. Dimana hasil rekomendasi parfumnya terdiri dari bunga kenanga, bunga lemon dan lavender untuk nuansa “bunga” dan kombinasi jeruk, jeruk mandarin, dan jeruk kalabrian untuk nuansa “segar”. Padahal nilai *cosine* LLM tidak terlalu tinggi jika dibandingkan dengan model Word2Vec. Rata - rata skor *cosine similarity* setiap model ditunjukkan pada Tabel 4. Rerata skor diambil dari penjumlahan seluruh nilai *cosine similarity* hasil rekomendasi dibagi top-3 hasil rekomendasi.

Tabel 4. Nilai *Cosine Similarity* Rata-rata Tiap Model

<i>Input</i>	Metode/Model	Rerata Nilai <i>Cosine Similarity</i>
Saya ingin parfum yang tercium wangi bunga namun segar	TF-IDF	0,14
	Word2Vec	0,96
	LLM	0,68

Berdasarkan Tabel 4 diatas, meskipun nilai skor *cosine similarity* *Word2Vec* memiliki angka yang paling tinggi namun LLM memberikan hasil rekomendasi yang lebih semantik dan kontekstual.

### 3.3. Analisis Kualitatif – Berdasarkan Responden

Metode evaluasi telah dilakukan dengan melibatkan 3 responden yang sehari-hari aktif menggunakan parfum atau diantaranya merupakan kolektor parfum. Responden diminta memasukkan pertanyaan “rekomendasikan parfum untuk olahraga”. Hasil Rekomendasi ditunjukkan pada Tabel 5. Setelah sistem menampilkan hasilnya, responden diminta memberi skor relevansi dari skala 0 hingga 1 untuk setiap rekomendasi. Tujuan evaluasi ini adalah menilai persepsi relevansi dari hasil sistem, bukan akurasi teknis. *Precision Matrix* berdasarkan responden ditunjukkan pada Tabel 6.

Tabel 5. Hasil Rekomendasi Berdasarkan Input

<i>Input</i>	Metode/Model	Rekomendasi 1 (Nama Parfum   Notes)	Rekomendasi 2 (Nama Parfum   Notes)	Rekomendasi 3 (Nama Parfum   Notes)
	TF-IDF	Montale Leather Patchouli Eau de	Frapin The Orchid Man Eau de Parfum	Maitre Parfumeur et

rekomendasikan parfum untuk olahraga		Parfume   wangi kayu; balsam labdanum; kayu oud; nilam; kayu cedar	jeruk bergamot; lada hitam; melati; wangi kulit; ambergris/musk	Gantier Santal Noble Eau de Parfume   biji kopi; minyak rempah-rempah; kayu cendana
	Word2Vec	Juliette Has a Gun Not a Perfume Eau de Parfum   sweet ambroxan; sweet amber cetalox	What We Do Is Secret Monoscent G Eau de Parfum   Sweet Galaxolide Super	Monyette Paris Eau de Parfum Spray   Bunga Kacapiring ( <i>gardenia</i> ); bunga Muguet Perancis; bunga kenanga
	LLM	J.F. Schwarzlose Zeitgeist Eau de Parfum   Amber balsam; balsam Peru; rumput laut; jeruk	Ensar Oud - American Oud   coklat bakar; wangi kulit; wangi resin; kayu oud	Parfums de Nicolai - Cologne Sologne Eau de Cologne   Neroli; kayu cendana; buah anggur; buah lemon

Tabel 6. Hasil Evaluasi Kualitatif Responden Berdasarkan Hasil Rekomendasi (R)

Metode	R1	R2	R3	Rata-rata
TF-IDF	0	0	0	0
Word2Vec	0	0	1	0,34
LLM	1	0	1	0,67

Berdasarkan inputan “rekomendasikan parfum untuk olahraga” pada Tabel 6, responden berasumsi bahwa secara umum parfum untuk olahraga harus segar atau segar dengan nuansa bunga. Dapat diuraikan dalam metrik Presisi (*precision*) dengan rumus sebagai berikut :

$$Precision@K = \frac{\text{jumlah parfum yang relevan (menurut responden)}}{\text{jumlah parfum yang direkomendasikan (k)}}$$

Maka presisi (*precision*) TF-IDF, Word2Vec dan LLM dapat dihitung sebagai berikut :

$$Presisi TF - IDF = \frac{0}{3} = 0$$

$$Presisi Word2Vec = \frac{1}{3} = 0,34$$

$$Presisi TF - IDF = \frac{2}{3} = 0,67$$

Berdasarkan evaluasi menggunakan metrik *precision@3* yang melibatkan 3 responden, diperoleh hasil bahwa: metode TF-IDF menghasilkan *precision* sebesar 0.00, Word2Vec sebesar 0.34, dan LLM sebesar 0.67. Nilai *precision* yang rendah pada TF-IDF menunjukkan bahwa pencocokan berbasis frekuensi kata tidak cukup representatif untuk memahami karakteristik aroma yang diinginkan pengguna. Sementara itu, Word2Vec memperlihatkan performa lebih baik karena

mampu menangkap kesamaan semantik antar kata. Namun demikian, hasil terbaik diperoleh dari metode LLM yang mampu memahami konteks kalimat dan makna deskriptif secara lebih utuh, sehingga menghasilkan rekomendasi yang lebih relevan secara subjektif menurut persepsi pengguna.

Berdasarkan hasil evaluasi kuantitatif dan kualitatif, pendekatan menggunakan *Large Language Model* (LLM), khususnya model *Sentence Transformer* all-MiniLM-L6-v2, menunjukkan performa yang unggul dalam konteks sistem rekomendasi parfum berbasis teks deskriptif. Meskipun skor *cosine similarity* dari metode Word2Vec cenderung lebih tinggi secara numerik, namun hasil rekomendasi yang diberikan oleh LLM dinilai lebih relevan secara semantik oleh para responden.

Hal ini menunjukkan bahwa LLM mampu memahami konteks dan makna kalimat secara menyeluruh, tidak hanya berdasarkan frekuensi atau kedekatan vektor kata. Dalam lingkup seperti parfum, di mana deskripsi dan aroma seringkali bersifat subjektif, metaforis, dan naratif (misalnya: "aroma lembut seperti pagi musim semi"), pendekatan model yang dapat teks secara lebih mendalam menjadi sangat penting. LLM dapat menangkap nuansa deskriptif yang kompleks tersebut, sehingga menghasilkan rekomendasi yang lebih personal dan sesuai ekspektasi pengguna.

Selain itu, keunggulan lain dari LLM adalah kemampuannya untuk melakukan generalisasi terhadap deskripsi baru yang belum pernah muncul dalam data pelatihan, berkat proses *pretraining* yang luas pada berbagai jenis teks. Ini memberikan fleksibilitas dalam skenario dunia nyata, di mana pengguna bisa memasukkan deskripsi yang unik, bebas, dan personal seperti inputan yang telah diuji dan uraikan.

#### 4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem rekomendasi parfum berbasis teks deskriptif dengan pendekatan *Large Language Model* (LLM), khususnya menggunakan arsitektur *Sentence Transformer* all-MiniLM-L6-v2. Sistem ini mampu memahami konteks kolom dataset deskriptif dan notes aroma yang bersifat subjektif dan naratif, seperti notes aroma "segar", "lembut", "maskulin", "woody" atau "balsamic" yang sulit ditangkap oleh pendekatan berbasis statistik biasa.

Hasil eksperimen menunjukkan bahwa meskipun metode Word2Vec menghasilkan skor *cosine similarity* yang lebih tinggi secara numerik, model LLM justru memberikan hasil rekomendasi yang paling relevan berdasarkan persepsi pengguna. Evaluasi kualitatif oleh responden juga mendukung bahwa LLM lebih unggul dalam menangkap makna semantik dari deskripsi yang diberikan. Hal ini didukung dengan nilai *precision* untuk LLM sebesar 0.67. Perbandingan antara model TF-IDF, Word2Vec, dan LLM yang telah dibentuk memperkuat bahwa kemampuan LLM dalam memahami keseluruhan kalimat dan konteks menjadikannya pendekatan yang lebih unggul untuk sistem rekomendasi berbasis teks, terutama pada domain yang bersifat deskriptif dan emosional seperti parfum.

Ke depannya, penelitian ini dapat dikembangkan dengan memperluas jumlah data, melakukan *fine-tuning* model LLM lebih jauh dengan membandingkan arsitektur LLM lainnya seperti *RoBERTa-based Models* atau *DistilBERT-based Models* pada korpus lokal (bahasa Indonesia), serta mengintegrasikan komponen lain seperti ulasan pengguna atau preferensi historis untuk meningkatkan personalisasi sistem rekomendasi.

#### DAFTAR PUSTAKA

- [1] B. C. L. Rodrigues, V. V. Santana, L. P. Queiroz, C. M. Rebello, and I. B. R. Nogueira, "Harnessing graph neural networks to craft fragrances based on consumer feedback," *Comput Chem Eng*, vol. 185, no. March, p. 108674, 2024, doi: 10.1016/j.compchemeng.2024.108674.
- [2] M. I. Z. K. B. Sienatra, "Analisis Preferensi Konsumen Memilih Produk Parfum Lembrou," *Parsimonia: Jurnal Akuntansi, Manajemen dan Bisnis*, vol. 11, no. 2, pp. 66-86, 2024.

- [3] R. Lutan and C. Badica, "Personalized Quiz-Based Perfume Recommender System Using Social Data," 2023, pp. 30–43. doi: 10.1007/978-3-031-41774-0\_3.
- [4] H. K. Omar, M. Frikha, and A. K. Jumaa, "Improving Big Data Recommendation System Performance Using NLP Techniques With Multi-attributes," *Informatica (Slovenia)*, vol. 48, no. 5, pp. 63–70, Feb. 2024, doi: 10.31449/inf.v48i5.5255.
- [5] T. Xu and B. Li, "KELLM: Knowledge-Enhanced Label-Wise Large Language Model for Safe and Interpretable Drug Recommendation," *Electronics (Switzerland)*, vol. 14, no. 1, Jan. 2025, doi: 10.3390/electronics14010154.
- [6] L. Zhang, Y. Hu, W. Li, Q. Bai, and P. Nand, "LLM-AIDSim: LLM-Enhanced Agent-Based Influence Diffusion Simulation in Social Networks," *Systems*, vol. 13, no. 1, Jan. 2025, doi: 10.3390/systems13010029.
- [7] W. S. Kim, S. Lim, G. W. Kim, and S. M. Choi, "Extracting Implicit User Preferences in Conversational Recommender Systems Using Large Language Models," *Mathematics*, vol. 13, no. 2, Jan. 2025, doi: 10.3390/math13020221.
- [8] Y. Tokutake and K. Okamoto, "Can Large Language Models Assess Serendipity in Recommender Systems?," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 28, no. 6, pp. 1263–1272, Nov. 2024, doi: 10.20965/jaciii.2024.p1263.
- [9] G. Hui, C. Mang, Z. LiQing, and X. ShiKun, "Research on Personalized Recommendation Algorithms Based on User Profile," *Int J Adv Comput Sci Appl*, vol. 15, no. 3, pp. 289–298, 2024.
- [10] R. Goodwin *et al.*, "AI for Fragrance Design," 2017, [Online]. Available: <http://www.basenotes.net/>,
- [11] J. Carles, *A Method of Creation In Parfumery*, 1st ed. 2006.
- [12] Adline Twince Tobing *et al.*, "Men's Perfume Recommendation System Using Analytic Hierarchy Process (AHP) and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) Method," *International Journal of Science, Technology & Management*, vol. 4, no. 6, pp. 1552–1559, 2023, doi: 10.46729/ijstm.v4i6.1002.
- [13] I. Ramadhani, A. Zakir, and A. M. Elhanafi, "Implementasi Metode Weighted Product Untuk Menentukan Alternatif Perfume Terbaik ( Studi Kasus : Quality Perfume )," *International Journal of Data Science and Visualization*, vol. 3, no. 1, pp. 1–11, 2024.
- [14] N. S. Agashe, "Product Recommender Chat Bot," *International Journal of Engineering Research & Technology (IJERT)*, vol. 10, no. 06, pp. 702–705, 2021.
- [15] K. Kalashi, S. Saed, and B. Teimourpour, "Sentiment-Driven Community Detection in a Network of Perfume Preferences," *ArXiv. Cornell University*, 2024.
- [16] S. Minaee *et al.*, "Large Language Models: A Survey," *arXiv:2402.06196v2 [cs.CL]*, 2024, [Online]. Available: <http://arxiv.org/abs/2402.06196>
- [17] A. Vaswani *et al.*, "Attention is all you need," *Adv Neural Inf Process Syst*, vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.
- [18] J. Achiam *et al.*, "GPT-4 Technical Report," *arXiv:2303.08774v6 [cs.CL] Cornell University*, vol. 4, pp. 1–100, 2023, [Online]. Available: <http://arxiv.org/abs/2303.08774>
- [19] J. Kim, K. Oh, and B. S. Oh, "An NLP-Based Perfume Note Estimation Based on Descriptive Sentences," *Applied Sciences (Switzerland)*, vol. 14, no. 20, 2024, doi: 10.3390/app14209293.
- [20] P. Chapman *et al.*, "CRISP-DM 1.0 Step-by-step data mining guide," DaimlerChrysler, 1999.