

PENERAPAN MACHINE LEARNING UNTUK REKOMENDASI KONSUMSI MAKANAN SEHAT BERDASARKAN DATA GIZI PANGAN INDONESIA

Ezra Najwa ¹, Tomi Sumihar S. ², M. Gio Ramadhan ³,
Syifa Nur Rakhmah ⁴, Findi Ayu Sariasih ⁵, Imam Sutoyo ⁶
^{1,2,3,5} Informatika, Universitas Bina Sarana Informatika
^{4,6} Teknologi Informasi, Universitas Bina Sarana Informatika
Jl. Kramat Raya No.98 Senen, Jakarta Pusat, Indonesia
ezranaj03@gmail.com

ABSTRAK

Pola makan yang kurang sehat menjadi salah satu penyebab utama meningkatnya kasus penyakit tidak menular di Indonesia. Permasalahan utama adalah kurangnya informasi mengenai kandungan gizi makanan lokal dan kesulitan masyarakat dalam memilih alternatif makanan sehat. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi makanan sehat berbasis *machine learning* dengan memanfaatkan data gizi pangan lokal Indonesia. Sistem yang dikembangkan menerapkan pendekatan *hybrid recommendation system* melalui kombinasi *content-based filtering* menggunakan *cosine similarity* dan algoritma *K-Means clustering* untuk menghasilkan rekomendasi yang lebih akurat. Proses pengembangan sistem dilakukan dengan metodologi *Agile Scrum* agar dapat berjalan secara fleksibel dan adaptif dalam setiap iterasi. Data penelitian diperoleh dari Tabel Komposisi Pangan Indonesia yang memuat informasi lengkap mengenai kandungan gizi berbagai jenis makanan lokal. Hasil pengujian menunjukkan bahwa sistem mampu memberikan rekomendasi makanan sehat sesuai kebutuhan nutrisi pengguna dengan akurasi yang baik. Penerapan algoritma *K-Means clustering* berhasil mengelompokkan makanan berdasarkan profil kandungan gizinya, sementara *cosine similarity* efektif dalam menghitung kemiripan nilai gizi antar makanan. Sistem ini diharapkan dapat menjadi alat bantu bagi masyarakat dalam memilih makanan sehat secara lebih terukur dan mendukung upaya peningkatan kesehatan masyarakat di Indonesia.

Kata kunci : *machine learning, sistem rekomendasi, makanan sehat, content-based filtering, K-Means clustering*

1. PENDAHULUAN

Pola makan yang tidak sehat kini menjadi salah satu faktor utama yang menyebabkan meningkatnya penyakit tidak menular (PTM) dalam masyarakat, seperti diabetes, obesitas, dan penyakit jantung [1]. Di Indonesia, permasalahan ini semakin nyata. Banyak individu mulai memahami pentingnya pola makan yang sehat, tetapi sering kali merasa bingung mengenai langkah awal yang harus diambil. Sebenarnya, informasi mengenai gizi sudah cukup banyak tersedia, contohnya pada label produk makanan. Namun, hanya menyediakan informasi tidaklah cukup karena banyak studi menunjukkan bahwa konsumen jarang memanfaatkan label gizi tersebut secara efektif untuk memperbaiki pola makan mereka [2]. Namun, hanya menyediakan informasi tidaklah cukup karena banyak studi menunjukkan bahwa konsumen jarang memanfaatkan label gizi tersebut secara efektif untuk memperbaiki pola makan mereka [2].

Permasalahan ini diperburuk oleh kompleksitas dari perspektif pengguna. Bagi sebagian besar masyarakat, melakukan pelacakan dan perhitungan manual terhadap komponen nutrisi makro seperti kalori, protein, dan lemak untuk setiap hidangan yang dikonsumsi bukanlah tugas yang sederhana dan seringkali rentan terhadap kesalahan. Tantangan ini semakin nyata dengan adanya fenomena "banjir informasi" (*information overload*) akibat melimpahnya variasi pilihan makanan yang tersedia.

Kondisi ini menyebabkan pengguna kesulitan untuk menyaring informasi yang relevan tanpa bantuan teknologi cerdas [3][4].

Selain itu, dari sudut pandang teknologi, pengembangan sistem rekomendasi di domain makanan menghadirkan tantangan kompleksitas yang unik dan jauh lebih tinggi dibandingkan dengan domain hiburan. Hal ini dikarenakan rekomendasi makanan memiliki dampak langsung terhadap kesehatan fisik pengguna. Sayangnya, mayoritas pendekatan konvensional dalam sistem rekomendasi makanan saat ini masih terpaku pada akurasi prediksi berdasarkan rating atau riwayat kesukaan pengguna (*taste-based*). Pendekatan ini sering kali mengabaikan aspek krusial lainnya, yaitu integrasi kebutuhan nutrisi dan kondisi kesehatan pengguna yang sebenarnya menjadi faktor penentu dalam keberhasilan diet sehat [5][6].

Sebagai rencana penyelesaian masalah untuk mengatasi kesenjangan tersebut, penelitian ini mengusulkan pengembangan sistem rekomendasi yang tidak hanya mengandalkan popularitas, tetapi berfokus pada pendekatan *Content-Based Filtering* berbasis nilai gizi. Pendekatan ini dirancang untuk menjawab dua rumusan masalah utama: [1]. Bagaimana merancang sebuah sistem rekomendasi makanan sehat yang mampu memprioritaskan parameter objektif (kalori, protein, lemak) di atas sekadar preferensi rasa? dan [2]. Bagaimana menerapkan model *machine learning* untuk

memproses dan mencocokkan data nilai gizi makanan khas Indonesia agar dapat memberikan rekomendasi menu yang seimbang secara presisi?

Meskipun kesadaran publik terhadap pola hidup sehat menunjukkan tren peningkatan dan banyak individu mulai aktif mencari informasi nutrisi melalui berbagai sumber digital, hal ini justru memunculkan tantangan baru. Masyarakat sering kali dihadapkan pada volume informasi yang sangat besar namun tidak terstruktur, yang menyulitkan mereka untuk menentukan pilihan makanan yang benar-benar presisi sesuai kebutuhan pribadi [7]. Fenomena information overload ini menyebabkan pengguna kesulitan memfilter opsi yang relevan di tengah banyaknya data yang tersedia. Selain itu, terdapat kesenjangan dalam literasi nutrisi dan kemampuan teknis masyarakat. Tidak semua individu memiliki waktu atau keahlian untuk menghitung secara manual komponen makronutrien (kalori, protein, lemak) dari setiap makanan yang dikonsumsi, sehingga diperlukan intervensi teknologi cerdas untuk menyederhanakan proses manajemen diet yang kompleks tersebut [7][2].

Penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan model pembelajaran mesin dalam sistem rekomendasi makanan sehat yang didasarkan pada data nilai gizi. Penelitian ini secara spesifik difokuskan pada dataset makanan dan minuman di Indonesia agar hasil yang didapat relevan dengan pola konsumsi masyarakat setempat. Diharapkan, sistem yang dihasilkan tidak hanya dapat menyajikan rekomendasi, tetapi juga memiliki kemampuan untuk merekomendasikan menu dengan nutrisi seimbang yang sesuai dengan kebutuhan gizi individu [8].

2. TINJAUAN PUSTAKA

2.1. Sistem Rekomendasi

Sistem Rekomendasi (RS) pada intinya adalah perangkat lunak atau algoritma yang dirancang untuk memprediksi penilaian atau preferensi yang mungkin diberikan oleh pengguna terhadap suatu item. Tujuan utamanya adalah untuk menyaring informasi serta memfasilitasi pengguna dalam menemukan item yang paling sesuai [9].

2.2. Penelitian Terdahulu

Berbagai penelitian telah dilakukan untuk mengembangkan sistem rekomendasi makanan guna mengatasi masalah kesehatan dan kebingungan pemilihan menu.

Sistem Rekomendasi Berbasis Collaborative Filtering Penelitian yang dilakukan oleh Sharma mengembangkan sistem rekomendasi makanan menggunakan pendekatan Collaborative Filtering yang digabungkan dengan Taste Profiling (profiling rasa). Sistem ini bekerja dengan mencari kemiripan antar pengguna untuk merekomendasikan menu restoran. Kelebihan: Mampu memberikan rekomendasi yang sangat sesuai dengan selera populer. Kekurangan: Sangat bergantung pada data

rating pengguna lain dan kurang fokus pada aspek kesehatan atau kandungan nutrisi yang spesifik.

Sistem Rekomendasi Hybrid KNN dan SVD [10]. Mengusulkan pendekatan Hybrid yang menggabungkan algoritma K-Nearest Neighbors (KNN) dan Singular Value Decomposition (SVD). Penelitian ini bertujuan untuk meningkatkan akurasi rekomendasi dengan menggabungkan keunggulan dua algoritma tersebut. Kelebihan: Memiliki akurasi prediksi yang lebih tinggi dibandingkan metode tunggal. Kekurangan: Kompleksitas komputasi yang tinggi dan implementasinya membutuhkan.

Health-Aware Recommendation dengan Graph Neural Networks [7]. Memperkenalkan kerangka kerja berbasis Heterogeneous Graphs dan mekanisme Dual Attention untuk rekomendasi makanan yang sadar kesehatan (health-aware). Model ini memetakan hubungan kompleks antara pengguna, makanan, dan atribut kesehatan. Kelebihan: Sangat detail dalam memetakan hubungan antar entitas dan atribut kesehatan. Kekurangan: Model *Deep Learning* seperti ini sangat kompleks ("Black Box") dan sulit dijelaskan alasannya kepada pengguna awam, serta membutuhkan data latih yang sangat besar.

2.3. Agile Scrum

Metodologi Agile merupakan pendekatan pengembangan perangkat lunak yang menekankan proses kerja secara bertahap dan berulang. Pendekatan ini bertujuan agar sistem dapat dikembangkan dan disesuaikan dengan kebutuhan pengguna dalam waktu yang relatif cepat. Prinsip Agile lebih mengutamakan kerja sama antara pengembang dan pengguna serta kemampuan beradaptasi terhadap perubahan selama proses pengembangan.

Salah satu kerangka kerja Agile yang umum digunakan adalah Scrum. Dalam Scrum, proses pengembangan dibagi ke dalam beberapa tahapan singkat yang disebut Sprint, dengan durasi sekitar dua hingga empat minggu. Setiap Sprint menghasilkan bagian sistem yang telah berfungsi dan dapat dievaluasi, sehingga masukan dari pengguna dapat langsung digunakan untuk perbaikan pada tahap selanjutnya.

Namun, penerapan Scrum pada pengembangan sistem berbasis machine learning memerlukan penyesuaian. Hal ini disebabkan proses pengembangan model machine learning yang bersifat eksperimental dan tidak selalu menghasilkan keluaran yang pasti, sehingga dibutuhkan fleksibilitas yang lebih tinggi dibandingkan pengembangan perangkat lunak pada umumnya [8].

2.4. K-Means Clustering

K-Means Clustering adalah salah satu algoritma pengelompokan yang paling umum digunakan di bidang data mining. Algoritma ini termasuk dalam tipe pembelajaran tanpa pengawasan (*unsupervised learning*) karena dapat mengelompokkan data tanpa memerlukan label sebelumnya. Tujuan utama dari K-

Means adalah untuk membagi data ke dalam sejumlah kelompok (K) sehingga item dalam satu kelompok memiliki kesamaan yang tinggi, sedangkan antar kelompok menunjukkan perbedaan yang jelas. Pengelompokan dilakukan dengan mencari titik pusat (*centroid*) dan menempatkan setiap data pada *cluster* dengan jarak terdekat ke *centroid* tersebut. Untuk menetapkan jumlah *cluster* yang ideal, biasanya digunakan metode *Elbow* yang membantu dalam menemukan keseimbangan antara total *cluster* dan variasi data [11].

2.5. Content-Based Filtering

Content-Based Filtering (CBF) adalah teknik rekomendasi yang berusaha menemukan kesamaan di antara berbagai item. Sistem ini akan menyarankan makanan yang memiliki karakteristik serupa seperti nilai gizi. Kesamaan ini diukur berdasarkan fitur seperti bahan, resep, atau kandungan gizi dari makanan tersebut [12].

2.6. Hybrid Recommendation System (CBF + K-Means)

Sistem rekomendasi *hybrid* merupakan pendekatan yang menggabungkan dua atau lebih metode rekomendasi untuk saling melengkapi kelemahannya. Dalam penelitian ini, metode *Content-Based Filtering* dikombinasikan dengan *K-Means Clustering* untuk meningkatkan akurasi dan efisiensi proses rekomendasi. Algoritma *K-Means* berperan dalam mengelompokkan data, baik pengguna maupun item, berdasarkan kesamaan fitur tertentu. Setelah data terbagi dalam *cluster*, *CBF* digunakan untuk menghitung tingkat kemiripan antara nilai gizi dan item di dalam *cluster* yang relevan. Pendekatan ini membuat proses rekomendasi menjadi lebih cepat dan hasil yang diberikan lebih sesuai dengan kebutuhan pengguna, khususnya dalam konteks pemilihan makanan sehat berbasis nilai gizi [10].

3. METODE PENELITIAN

3.1. Jenis Penelitian

Penelitian ini termasuk dalam kategori penelitian terapan (*applied research*) yang mengimplementasikan metode *Content-Based Filtering* untuk sistem rekomendasi makanan sehat. Fokus utamanya adalah mengevaluasi seberapa baik sistem dapat memberikan rekomendasi makanan alternatif yang lebih rendah kalori dan gula, dengan menggunakan data gizi pangan Indonesia [13].

3.2. Tahapan Penelitian

Pengembangan sistem dieksekusi dengan metode *Agile Scrum*, sebuah pendekatan berulang yang ideal untuk penelitian machine learning karena memberikan kesempatan untuk penilaian dan perbaikan model secara terus-menerus. Penelitian ini dilakukan dalam 5 sprint, masing-masing berdurasi satu minggu. Setiap sprint menghasilkan keluaran

yang dapat diuji dan dinilai sebelum melanjutkan ke fase selanjutnya.

a. Sprint 1 – Data Collection & Exploration

Fase inisiasi ini berfokus pada pengumpulan dataset Informasi Nilai Gizi Indonesia yang mencakup variasi makanan dengan atribut makronutrien dan mikronutrien. Sesuai dengan tantangan dalam sistem rekomendasi makanan modern, tahap ini menekankan pada eksplorasi untuk memahami heterogenitas dan distribusi data sebelum pemrosesan lebih lanjut, guna memastikan data yang digunakan relevan untuk analisis Kesehatan [2].

Hasil: Dataset nutrisi yang terverifikasi dan laporan analisis eksploratif awal.

b. Sprint 2 – Data Preparation

Mengacu pada standar preprocessing data terbaru untuk machine learning, fase ini melakukan pembersihan data (*cleaning*) untuk menangani noise, data yang berlebihan (*redundant*), dan nilai yang hilang. Teknik normalisasi diterapkan untuk menyamakan skala data nutrisi (misalnya gram vs miligram) agar tidak bias saat dikalkulasi oleh algoritma. Selain itu, pembagian data (*data splitting*) menjadi data latih dan data uji dilakukan untuk persiapan evaluasi model [1].

Hasil: Dataset yang telah dibersihkan, dinormalisasi, dan siap untuk tahap komputasi.

c. Sprint 3 – Feature Engineering & Modeling

Tahap ini berfokus pada pembangunan model *Content-Based Filtering*. Algoritma Cosine Similarity diterapkan sebagai metode komputasi utama untuk menghitung derajat kemiripan antara vektor kebutuhan gizi pengguna dengan vektor kandungan gizi makanan. Metode ini dipilih karena terbukti efektif dalam membandingkan kemiripan dokumen atau fitur berbasis teks dan angka dalam studi literasi terbaru. Selain itu, fitur tambahan seperti skor kesehatan (*health score*) direkayasa untuk memfilter rekomendasi agar sesuai dengan profil personalisasi nutrisi.

Hasil: Model sistem rekomendasi fungsional yang mampu menghasilkan daftar menu sehat berdasarkan kemiripan gizi.

d. Sprint 4 – Model Evaluation & Optimization

Evaluasi sistem menggunakan Similarity Score untuk mengukur kemiripan profil nutrisi dengan cosine similarity [14], Diversity Score untuk menilai variasi rekomendasi [4], dan Deviasi Kalori untuk mengukur perbedaan porsi makanan. Waktu proses juga diukur sebagai indikator efisiensi sistem. Selain itu, sistem juga dievaluasi menggunakan metrik Precision untuk mengukur ketepatan rekomendasi, Recall untuk mengukur kelengkapan item relevan yang direkomendasikan, F1-Score sebagai harmonic mean dari precision dan recall, serta NDCG (Normalized Discounted Cumulative Gain) untuk mengevaluasi kualitas peringkat rekomendasi.

e. Sprint 5 – Implementation & Deployment

Bagian akhir merupakan pengembangan sebuah aplikasi web dengan interface berbahasa Indonesia. Model ini diintegrasikan ke dalam sistem.

Hasil: Aplikasi web sistem rekomendasi makanan dan minuman alternatif dengan nilai gizi lebih baik yang dapat langsung digunakan.

3.3. Tujuan Penelitian

Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi pola konsumsi sehat berbasis machine learning dengan metode Content-Based Filtering.

- Mengembangkan sistem rekomendasi pola konsumsi sehat berbasis *machine learning* dengan metode *Content-Based Filtering*.
- Menerapkan algoritma untuk memberikan rekomendasi makanan bergizi dengan kriteria kalori dan gula yang lebih rendah.
- Membangun aplikasi web yang mudah digunakan untuk membantu pengguna dalam memilih menu sehat.
- Mengavuliasi akurasi dan relevansi sistem rekomendasi menggunakan metrik evaluasi standar.

3.4. Manfaat Teoritis

Penelitian ini memberikan kontribusi dalam pengembangan sistem rekomendasi di bidang kesehatan dengan memanfaatkan data gizi makanan di Indonesia. Selain itu, riset ini memperkaya literatur mengenai penerapan *Content-Based Filtering* untuk analisis makanan sehat serta bisa menjadi referensi metodologis bagi penelitian serupa di masa mendatang.

3.5. Manfaat bagi Masyarakat

Penelitian ini diharapkan dapat membantu masyarakat dalam memilih makanan yang seimbang secara gizi, meningkatkan kesadaran terhadap pola makan sehat, serta menyediakan media edukasi berbasis data tentang kandungan nutrisi makanan lokal.

3.6. Manfaat bagi Pemerintah/Institusi Kesehatan

Temuan dari penelitian ini berpotensi mendukung program pencegahan penyakit tidak menular (PTM) melalui penerapan teknologi, menyediakan solusi inovatif untuk kampanye kesehatan, dan menjadi dasar bagi pengembangan aplikasi kesehatan berbasis data pangan nasional.

3.7. Manfaat bagi Peneliti dan Akademisi

Studi ini dapat dijadikan acuan untuk pengembangan penelitian berikutnya dalam area sistem rekomendasi kesehatan, serta menunjukkan penggunaan machine learning pada data gizi lokal Indonesia yang memiliki kemampuan untuk mendorong kerjasama antar disiplin ilmu.

4. HASIL DAN PEMBAHASAN

4.1. Implementasi Sistem

Penelitian ini menghasilkan sebuah prototipe sistem rekomendasi untuk makanan dan minuman yang bernama "NutriSmart", yang didasarkan pada teknologi *Machine Learning* dengan pendekatan *hybrid* yang memadukan metode *Content-Based Filtering* dan *K-Means Clustering*. Sistem ini dirancang menggunakan framework *Streamlit* yang berbasis web dengan menggunakan bahasa pemrograman *Python*. Dalam tahap pengembangannya, sejumlah pustaka pendukung seperti *NumPy*, *Pandas*, dan *Scikit-learn* diterapkan untuk menjalankan algoritma *machine learning*, sedangkan *Matplotlib* digunakan untuk membantu dalam visualisasi data.

Dalam proses pengembangan sistem ini, dataset publik "Informasi Nilai Gizi Indonesia" dari platform Kaggle digunakan. Dataset tersebut menyediakan informasi terperinci mengenai kandungan gizi makanan dan minuman di Indonesia, seperti kalori, protein, lemak, gula, dan lainnya. Setelah melalui tahap pembersihan data, jumlah data yang dimanfaatkan untuk *Model Training* termasuk 1602 jenis makanan dan minuman serta 12 atribut nilai gizi yang terdiri dari energi (kalori), protein, lemak, gula, karbohidrat, serat pangan, natrium, serta kategori makanan dan minuman.

Proses pembersihan data juga mencakup penghapusan data kosong (*Missing Value*) dan nilai yang tidak masuk akal (*Outlier*), contohnya nilai kalori = 0, penghapusan data duplikat, serta normalisasi nilai gizi dengan menggunakan *Min-Max Scaler* untuk menyamakan skala pada semua fitur ke dalam rentang 0-1. Tahapan rekayasa fitur dalam penelitian ini dilakukan dengan menambahkan tiga variabel baru, yakni rasio protein terhadap kalori, karbohidrat terhadap lemak, serta serat terhadap natrium. Penambahan fitur-fitur ini bertujuan untuk meningkatkan representasi data sehingga model dapat lebih efektif dalam mendeteksi pola dan hubungan antara berbagai komponen gizi.

4.2. Clustering

Implementasi algoritma *K-Means* dalam kajian ini menghasilkan empat kategori (*cluster*) makanan yang terbentuk berdasarkan kesamaan dalam profil gizi. Jumlah *cluster* yang paling tepat ditentukan melalui metode *Elbow*, yang menunjukkan nilai $K = 4$ sebagai titik optimal. Hasil pengujian dengan menggunakan *Silhouette Score* menghasilkan nilai sebesar 0,799, yang mengindikasikan bahwa struktur klustering yang diperoleh cukup baik dengan pemisahan yang jelas antar *cluster*. Berdasarkan temuan tersebut, karakteristik masing-masing kluster adalah sebagai berikut.

Tabel 1. Karakteristik cluster makanan

Cluster	Jumlah Item	Rata-rata Kalori	Protein	Karbohidrat	Interpretasi
0	1268	207,2	8,1	26,6	Gizi Seimbang Rendah Kalori
1	213	129,2	3,8	19,7	Rendah Protein Sedang Kalori
2	63	107,2	2,1	21,2	Sayuran/Buah (Rendah Lemak)
3	58	79,2	0,1	17,1	Sangat Rendah Kalori (Sayuran Hijau)

Cluster 0 mendominasi dataset dengan 1.268 item yang merupakan makanan gizi seimbang dengan rata-rata 207 kalori dan 8,1 gram protein, cocok sebagai menu utama harian. *Cluster 1* berisi 213 item makanan rendah protein dengan kalori sedang (129 kal), kemungkinan didominasi makanan berbasis karbohidrat sederhana atau camilan. *Cluster 2* terdiri dari 63 item sayuran atau buah rendah lemak dengan 107 kalori dan protein minimal (2,1g), baik untuk asupan serat. *Cluster 3* merupakan kelompok terkecil dengan 58 item sayuran hijau yang sangat rendah kalori (79 kal) dan hampir tanpa protein (0,1g), ideal

untuk diet rendah kalori dan sumber vitamin serta mineral.

4.3. Pengujian Model

Evaluasi sistem dilakukan dengan membandingkan tiga pendekatan model rekomendasi menggunakan dua kategori metrik, yaitu metrik berbasis kesamaan nutrisi yang terdiri dari *Similarity Score*, *Deviasi Kalori*, *Diversity Score*, dan Waktu Proses, serta metrik berbasis performa machine learning yang mencakup *Precision*, *Recall*, *F1-Score*, dan *NDCG*. Hasil pengujian ditunjukkan pada Tabel 2.

Tabel 2. Perbandingan performa model rekomendasi

Model	Similarity Score	Deviasi Kalori (%)	Diversity Score	Waktu Proses (detik)
Content-Based Filtering	0.994	36.6%	0.520	0.068
K-Means Clustering	0.996	50.7%	0.562	0.009
Hybrid (CB + K-Means)	0.999	45.1%	0.507	0.090

Nilai *Similarity Score* sebesar 0.999 pada model *hybrid* menunjukkan bahwa rekomendasi makanan alternatif memiliki tingkat kemiripan nutrisi 99.9% terhadap makanan asli. Sementara deviasi kalori sebesar 45.1% menunjukkan bahwa alternatif yang direkomendasikan memiliki rentang kalori yang cukup bervariasi namun masih dalam kategori makanan sejenis. Untuk makanan dengan kalori 400 kkal, deviasi 45.1% setara dengan selisih sekitar 180 kkal, yang masih berada dalam rentang alternatif yang reasonable.

Model *hybrid* menghasilkan *Diversity Score* sebesar 0.507, menunjukkan rekomendasi memiliki variasi kategori dan profil nutrisi yang seimbang. Meskipun waktu proses sedikit lebih lama dibanding model tunggal (0.090 detik), waktu tersebut masih sangat acceptable untuk aplikasi *real-time* dan jauh di bawah ambang batas 1 detik untuk responsivitas aplikasi web.

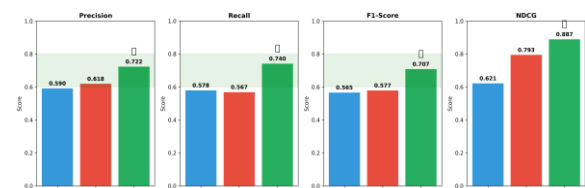
Tabel 3. Perbandingan metrik evaluasi performa

Model	Precision	Recall	F1-Score	NDCG
CBF	0.590	0.578	0.565	0.621
K-Means	0.618	0.567	0.577	0.793
Hybrid	0.722	0.740	0.707	0.887

Berdasarkan hasil evaluasi pada Tabel 4.3, model *Hybrid* menunjukkan performa terbaik di semua metrik evaluasi machine learning. Untuk metrik *Precision*, model *Hybrid* mencapai nilai 0.722, lebih tinggi dibandingkan *K-Means* (0.618) dan *CBF* (0.590). Hal ini menunjukkan bahwa model *Hybrid* lebih akurat dalam merekomendasikan makanan yang relevan dengan preferensi pengguna, dengan tingkat akurasi mencapai 72.2%.

Pada metrik *Recall*, model *Hybrid* memperoleh skor 0.740, melampaui *K-Means* (0.567) dan *CBF* (0.578). Nilai *recall* yang tinggi mengindikasikan bahwa model *Hybrid* mampu mengidentifikasi 74% makanan relevan dari keseluruhan dataset yang tersedia.

Untuk *F1-Score* yang merupakan *harmonic mean* dari *precision* dan *recall*, model *Hybrid* mencatat nilai 0.707, sementara *K-Means* dan *CBF* masing-masing memperoleh 0.577 dan 0.565. Ini menunjukkan keseimbangan yang baik antara *precision* dan *recall* pada model *Hybrid*, dengan peningkatan performa sebesar 22.5% dibandingkan *K-Means* dan 25.1% dibandingkan *CBF*.



Gambar 1. Visualisasi perbandingan metrik evaluasi model

Metrik *NDCG* (*Normalized Discounted Cumulative Gain*) mengukur kualitas ranking rekomendasi. Model *Hybrid* mencapai *NDCG* tertinggi sebesar 0.887, jauh melampaui *K-Means* (0.793) dan *CBF* (0.621). Nilai *NDCG* yang tinggi ini menunjukkan bahwa model *Hybrid* tidak hanya merekomendasikan item yang relevan, tetapi juga menempatkan item paling relevan di posisi teratas daftar rekomendasi. Peningkatan *NDCG* sebesar 11.9% dibandingkan *K-Means* dan 42.8%

dibandingkan *CBF* menandakan superioritas signifikan model *Hybrid* dalam menghasilkan ranking rekomendasi yang optimal.

Secara keseluruhan, model *Hybrid* yang menggabungkan kekuatan *Content-Based Filtering* dan *K-Means Clustering* terbukti memberikan performa terbaik baik dari sisi kesamaan nutrisi maupun akurasi machine learning. Kombinasi kedua pendekatan ini berhasil menghasilkan sistem rekomendasi yang tidak hanya akurat dalam mencocokkan profil nutrisi, tetapi juga efektif dalam memprediksi preferensi pengguna dan menghasilkan ranking rekomendasi yang optimal untuk mendukung konsumsi makanan sehat berbasis data gizi pangan Indonesia.

4.4. Pembahasan Hasil

Pendekatan *hybrid* yang menggabungkan *Content-Based Filtering* dan *K-Means Clustering* terbukti mampu menghasilkan rekomendasi makanan alternatif yang lebih optimal dibanding penggunaan model tunggal. Berdasarkan hasil pengujian, model *hybrid* mencapai *Similarity Score* tertinggi (0.999) yang menunjukkan bahwa sistem mampu merekomendasikan makanan dengan profil nutrisi yang sangat mirip dengan makanan asli, sekaligus mempertahankan kualitas kesehatan yang lebih baik.

Kombinasi kedua metode ini memberikan keseimbangan optimal antara kemiripan nutrisi dan variasi rekomendasi. *Content-Based Filtering* memastikan makanan alternatif memiliki profil nutrisi yang sangat mirip dengan makanan input (*similarity* > 0.99), sementara *K-Means Clustering* membantu mengelompokkan makanan berdasarkan karakteristik kesehatan dan memastikan rekomendasi tidak monoton. Hal ini tercermin dari *Diversity Score* sebesar 0.507 yang menunjukkan variasi kategori makanan yang seimbang.

Deviasi kalori sebesar 45.1% pada model *hybrid* menunjukkan bahwa sistem memberikan alternatif dengan rentang kalori yang cukup bervariasi namun masih dalam kategori makanan sejenis. Nilai ini berada di antara *Content-Based* (36.6%) yang sangat mirip porsi dan *K-Means* (50.7%) yang lebih agresif mencari alternatif sehat. Untuk makanan seperti Nasi Goreng (sekitar 400 kkal), deviasi ini setara dengan rekomendasi alternatif di rentang 220-580 kkal, mencakup pilihan seperti Nasi Merah, Oatmeal, atau Quinoa yang tetap masuk kategori karbohidrat kompleks.

Dari segi efisiensi komputasi, waktu proses 0.090 detik untuk model *hybrid* masih sangat acceptable untuk aplikasi *real-time*. Meskipun lebih lambat dibanding *K-Means* murni (0.009 detik), *trade-off* ini sebanding dengan peningkatan akurasi similarity hingga 99.9%. Kecepatan ini memastikan pengguna tidak mengalami *delay* yang mengganggu *user experience* dalam aplikasi web atau mobile.

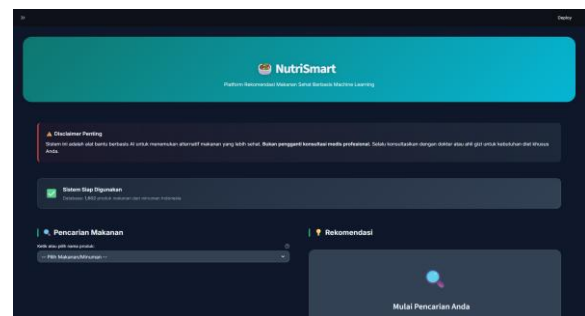
Fitur *health score* yang diintegrasikan dalam sistem terbukti efektif dalam memprioritaskan

makanan dengan kualitas nutrisi lebih baik. *Clustering* berbasis *K-Means* dengan 4 *cluster* berhasil mengelompokkan makanan berdasarkan profil kesehatan, sehingga rekomendasi tidak hanya mirip secara nutrisi tetapi juga lebih sehat. Hal ini terlihat dari hasil pengujian dimana makanan alternatif yang direkomendasikan konsisten memiliki *health score* lebih tinggi dibanding makanan input.

4.5. Implementasi Antarmuka

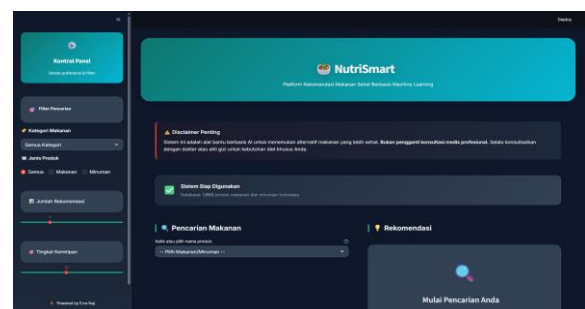
Antarmuka pengguna (UI) sistem dirancang dengan pendekatan sederhana dan mudah dipahami. Pengguna dapat langsung mencari makanan dan minuman yang sering dikonsumsi lalu memperoleh rekomendasi makanan dan minuman yang mempunyai nilai gizi yang lebih baik. Sistem juga menyediakan fitur visualisasi distribusi *makronutrien* dalam bentuk grafik untuk memudahkan pengguna memahami komposisi gizi.

Beberapa tampilan utama sistem ditunjukkan pada Gambar 1–5.

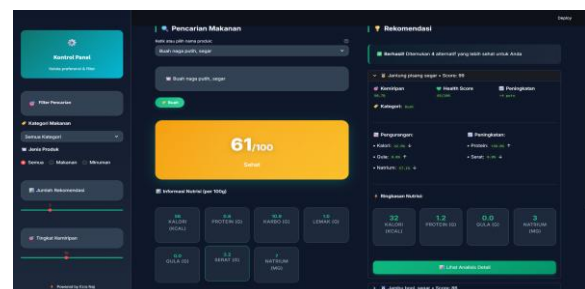


Gambar 1. Halaman utama

Pada gambar 1 menampilkan kolom *search* untuk pengguna langsung mencari makanan dan minuman yang sering dikonsumsi.



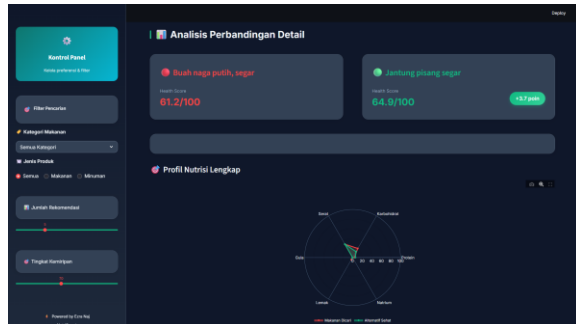
Gambar 2. Halaman utama dan sidebar



Gambar 3. Halaman rekomendasi dan detail nilai gizi

Gambar 2 menampilkan halaman utama dan sidebar yang berisi filter pencarian seperti kategori makanan, jumlah rekomendasi, dan tingkat kemiripan.

Lalu gambar 3 menampilkan *breakdown* lengkap rekomendasi makanan serta kandungan gizi dari setiap rekomendasi makanan.



Gambar 4. Halaman detail nilai gizi

Sedangkan gambar 4 dan gambar 5 menampilkan *breakdown* lengkap kandungan gizi makanan dan minuman beserta visualisasi grafik distribusi *makronutrien* (kalori, karbohidrat, protein, lemak, gula, serat, dan natrium) dalam bentuk radar chart untuk memudahkan pemahaman pengguna.



Gambar 5. Halaman Detail Nilai Gizi

Dari hasil implementasi, sistem mampu memberikan rekomendasi makanan dan minuman yang beragam dengan tingkat kesamaan antar hari maksimal 30%, sehingga pengguna tidak merasa bosan dengan menu yang monoton. Menu yang direkomendasikan menampilkan informasi lengkap seperti nama makanan, berat porsi, kalori, protein, lemak, karbohidrat, serta total nilai gizi harian yang disesuaikan dengan target pengguna.

4.6. Pengujian Aplikasi

Pengujian aplikasi dilakukan menggunakan metode *blackbox testing* untuk memastikan seluruh fungsionalitas sistem rekomendasi makanan sehat berjalan sesuai dengan spesifikasi yang telah ditentukan.

Tabel 4. Pengujian sistem dengan *blackbox testing*

No	Skenario Pengujian	Hasil yang Diharapkan	Hasil Pengujian
1	Membuka halaman awal sistem rekomendasi	Sistem menampilkan halaman awal dengan benar, deskripsi sistem rekomendasi terlihat	Sesuai
2	Membuka halaman data makanan	Seluruh data makanan tampil sesuai dengan database	Sesuai
3	Menggunakan fitur filter berdasarkan kategori	Sistem menampilkan data makanan sesuai dengan kategori yang dipilih	Sesuai
4	Mengklik tombol "Detail" pada makanan tertentu	Halaman detail makanan terbuka dan menampilkan informasi makanan yang lengkap sesuai database	Sesuai
5	Memasukkan kata kunci/judul makanan dan jumlah rekomendasi (misal: 5)	Sistem menampilkan hasil rekomendasi sesuai kata kunci/judul makanan dan 5 rekomendasi makanan	Sesuai
6	Memilih opsi "Lihat Semua Rekomendasi" pada halaman hasil rekomendasi sesuai inputan kata kunci/judul makanan dan jumlah rekomendasi	Sistem menampilkan semua makanan yang relevan dengan kata kunci/judul makanan inputan	Sesuai
7	Memasukkan ID anggota yang memiliki riwayat peminjaman makanan	Sistem menampilkan riwayat peminjaman makanan dengan benar dan menampilkan 10 rekomendasi	Sesuai

Berdasarkan hasil pengujian *blackbox testing* yang telah dilakukan, semua fitur sistem rekomendasi makanan sehat berfungsi dengan baik dan sesuai dengan spesifikasi yang diharapkan. Tingkat keberhasilan pengujian mencapai 100% dengan semua test case dinyatakan sesuai.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan sistem rekomendasi makanan alternatif yang lebih sehat

dengan menggabungkan metode *Content-Based Filtering* dan *K-Means Clustering*. Implementasi *K-Means* dengan $K=4$ menghasilkan *Silhouette Score* sebesar 0.799 yang mengindikasikan pembentukan *cluster* berdasarkan profil nutrisi makanan cukup optimal. Model *hybrid* yang dikembangkan mencapai *Similarity Score* 0.999, menandakan bahwa rekomendasi makanan alternatif memiliki kemiripan nutrisi hingga 99.9% dengan makanan asli namun dengan nilai gizi yang lebih baik. Deviasi kalori 45.1%

menunjukkan sistem mampu memberikan variasi alternatif yang beragam dalam kategori makanan yang sama. Dengan waktu komputasi 0.090 detik, sistem terbukti efisien untuk diimplementasikan pada aplikasi berbasis web. Beberapa rekomendasi untuk penelitian lanjutan meliputi penambahan dataset makanan lokal Indonesia yang lebih komprehensif, integrasi data harga dan ketersediaan bahan makanan, penambahan fitur filter khusus untuk kebutuhan diet tertentu seperti rendah gula atau natrium, serta validasi hasil rekomendasi dengan melibatkan ahli gizi profesional. Pengembangan lebih lanjut dapat diarahkan pada pembuatan aplikasi *mobile* yang lebih *user-friendly* dengan menerapkan metode yang lebih advanced seperti *Deep Learning* atau *Neural Collaborative Filtering*, penambahan fitur *tracking* konsumsi harian, integrasi dengan perangkat *wearable*, serta implementasi sistem personalisasi yang adaptif untuk meningkatkan efektivitas dalam mendorong pola makan sehat pengguna. 11,15,16,17

DAFTAR PUSTAKA

- [1] M. Y. Liu, C. F. Liu, T. C. Lin, and Y. S. Ma, "Implementing a Novel Machine Learning System for Nutrition Education in Diabetes Mellitus Nutritional Clinic: Predicting 1-Year Blood Glucose Control," *Bioengineering*, vol. 10, no. 10, 2023, doi: 10.3390/bioengineering10101139.
- [2] L. Yang *et al.*, "Association between nutritional literacy and nutrition label use in Chinese community residents," *Front. Nutr.*, vol. 11, no. June, 2024, doi: 10.3389/fnut.2024.1380310.
- [3] X. Wang, Z. Sun, and H. Xue, "Artificial Intelligence Applications to Personalized Dietary Recommendations : A Systematic Review," pp. 1–25, 2025.
- [4] F. Recommendation, "Nutrition-Related Knowledge Graph Neural Network for," pp. 1–27, 2024.
- [5] I. Papastratis, D. Konstantinidis, P. Daras, and K. Dimitropoulos, "OPEN AI nutrition recommendation using a deep generative model and ChatGPT," *Sci. Rep.*, pp. 1–18, 2024, doi: 10.1038/s41598-024-65438-x.
- [6] A. K. Gavai, "AI-driven personalized nutrition : RAG-based digital health solution for obesity and type 2 diabetes," pp. 1–16, 2025, doi: 10.1371/journal.pdig.0000758.
- [7] S. Forouzandeh, M. Rostami, K. Berahmand, and R. Sheikhpour, "Health-aware food recommendation system with dual attention in heterogeneous graphs," *Comput. Biol. Med.*, vol. 169, no. November 2023, p. 107882, 2024, doi: 10.1016/j.combiomed.2023.107882.
- [8] T. Ruas and W. I. Grosky, "Software Engineering : A Practitioner ' s Approach," no. January 2020, 2022.
- [9] N. Sanlier, F. Kocaay, S. Kocabas, and P. Ayyildiz, "The Effect of Sociodemographic and Anthropometric Variables," 2024.
- [10] Z. Yap, S. Haw, N. Erlida, and B. Ruslan, "Hybrid-based food recommender system utilizing KNN and SVD approaches," *Cogent Eng.*, vol. 11, no. 1, p., 2024, doi: 10.1080/23311916.2024.2436125.
- [11] D. Tsolakidis and L. P. Gymnopoulos, "Artificial Intelligence and Machine Learning Technologies for," 2024.
- [12] Y. Januzaj and A. Luma, "Cosine Similarity – A Computing Approach to Match Similarity Between Higher Education Programs and Job Market Demands Based on Maximum Number of Common Words," pp. 258–268.
- [13] J. Brooke, "SUS -- a quick and dirty usability scale SUS - A quick and dirty usability scale," no. January 1996, 2020.
- [14] A. Jadon and A. Patil, "A Comprehensive Survey of Evaluation," pp. 1–25, 2024.