



Artikel Penelitian

Implementasi Machine Learning dalam Sistem Prediksi dan Rekomendasi Program Diet Terintegrasi LLM

Endah Septa Sintiya ^{a,*}, Sely Ruli Amanda ^b, Candra Bella Vista ^c, Agung Nugroho Pramudhita ^d

^a Program Studi D2 Pengembangan Perangkat Lunak Situs, Jurusan Teknologi Informasi, Politeknik Negeri Malang, Jl. Soekarno Hatta No.9, Jatimulyo, Kota Malang, Indonesia

^{b,c,d} Program Studi D4 Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Malang, Jl. Soekarno Hatta No.9, Jatimulyo, Kota Malang, Indonesia

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 07 Juli 2025

Revisi Akhir: 02 September 2025

Diterbitkan Online: 07 September 2025

KATA KUNCI

Machine learning,
Prediksi diet,
Sistem Rekomendasi,
Program diet,
Integrasi LLM

KORESPONDENSI

E-mail: e.septa@polinema.ac.id *

A B S T R A C T

Masalah malnutrisi, baik dalam bentuk kelebihan maupun kekurangan berat badan, masih menjadi tantangan kesehatan global. Gaya hidup urban yang tidak sehat dan keterbatasan akses terhadap intervensi gizi yang tepat turut memperparah permasalahan ini. Pendekatan berbasis teknologi seperti *machine learning* dan *Large Language Models* (LLM) menawarkan peluang untuk meningkatkan efektivitas pengelolaan pola makan. Penelitian ini mengusulkan pengembangan sistem prediksi dan rekomendasi program diet berbasis *machine learning* dan terintegrasi LLM yang diterapkan pada Cafe NUT Castle. Sistem dikembangkan untuk mendigitalisasi pencatatan data komposisi tubuh, memprediksi program diet (penurunan berat badan, peningkatan berat badan, dan penurunan lemak tubuh) menggunakan algoritma Random Forest, serta menghasilkan rekomendasi diet awal yang dipersonalisasi melalui integrasi Gemini Flash-Lite API. Berdasarkan hasil pengujian, model prediksi mencapai akurasi sebesar 93% pada data uji dan 84% pada 50 data baru. Evaluasi terhadap rekomendasi diet yang dihasilkan oleh LLM menunjukkan tingkat kelayakan sebesar 86,6% yang dikategorikan sebagai sangat layak. Hasil ini menunjukkan bahwa sistem yang dikembangkan tidak hanya akurat dalam memprediksi program diet, tetapi juga efektif dalam memberikan rekomendasi awal yang dapat mendukung pengambilan keputusan dalam layanan konsultasi gizi secara digital.

1. PENDAHULUAN

Masalah malnutrisi, mencakup kondisi kelebihan berat badan maupun kekurangan berat badan merupakan tantangan kesehatan global yang signifikan. Data World Health Organization (WHO) pada tahun 2022 menunjukkan bahwa sebanyak 2,5 miliar orang dewasa mengalami kelebihan berat badan dan 390 juta lainnya mengalami kekurangan berat badan [1]. Masalah ini diperparah oleh gaya hidup urban yang cenderung tidak sehat, termasuk konsumsi makanan ultra proses yang tinggi [2]. Berbagai penelitian telah menunjukkan korelasi antara konsumsi makanan ultra proses dengan kualitas diet yang lebih rendah dan peningkatan risiko obesitas [3], [4]. Untuk mengatasi masalah ini, pemanfaatan teknologi seperti *machine learning* (ML) telah

banyak dieksplorasi. Penelitian sebelumnya menunjukkan efektivitas algoritma seperti *Random Forest* dalam memprediksi obesitas dan penyakit terkait dengan akurasi yang tinggi [5], [6], [7]. Studi lain juga membuktikan bahwa ML dapat menjadi alat yang efektif untuk mengidentifikasi risiko malnutrisi [8] dan memprediksi kategori berat badan dengan akurasi mencapai 92% dan 98% [9], [10]. Namun, sebagian besar penelitian tersebut cenderung berhenti pada tahap prediksi atau klasifikasi tanpa melangkah lebih jauh untuk pemberian rekomendasi tindakan yang terpersonalisasi.

Di sisi lain, studi kasus pada Cafe NUT Castle di Kota Malang menemukan beberapa tantangan operasional. Proses pencatatan data komposisi tubuh pelanggan masih dilakukan secara manual menggunakan media kertas, yang rentan terhadap risiko

kehilangan data, kerusakan, dan kesulitan dalam analisis data historis. Selain itu, proses penentuan program diet yang dilakukan secara manual terkadang dapat menimbulkan kesalahan, terutama jika ahli gizi tidak berada di tempat.

Untuk menjembatani kesenjangan tersebut, penelitian ini mengembangkan sebuah sistem terintegrasi berbasis web. Sistem ini tidak hanya menerapkan *machine learning* dengan algoritma *Random Forest* untuk memprediksi program diet secara akurat, tetapi juga mengintegrasikan hasilnya dengan *Large Language Models* (LLM) untuk menghasilkan rekomendasi diet awal yang dipersonalisasi. Dengan demikian, penelitian ini bertujuan untuk (1) mendigitalisasi pencatatan data untuk meningkatkan integritas dan aksesibilitas data, (2) membangun model prediksi program diet yang andal, dan (3) mengevaluasi kelayakan rekomendasi diet yang dihasilkan oleh LLM sebagai referensi awal bagi ahli gizi dan pelanggan.

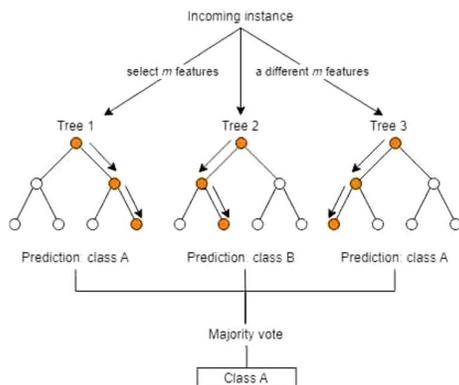
2. METODE

2.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data historis komposisi tubuh dari 369 pelanggan Cafe NUT Castle. Dataset terdiri dari 11 fitur prediktor (umur, tinggi badan, berat badan, lemak tubuh, lemak perut, kepadatan tulang, kebutuhan kalori, usia sel, massa otot, kadar air) dan 1 fitur target (program diet).

2.2. Algoritma Random Forest

Random Forest adalah metode *ensemble learning* yang tergolong dalam *supervised learning* dan digunakan untuk tugas klasifikasi maupun regresi. Algoritma ini bekerja dengan membangun sejumlah besar pohon keputusan (*decision trees*) pada berbagai sub-sampel dari dataset. Untuk klasifikasi, hasil akhir ditentukan melalui mekanisme pemungutan suara mayoritas (*majority voting*) dari semua pohon individu seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Ilustrasi Mekanisme Kerja Algoritma Random Forest

Ilustrasi tersebut menunjukkan bagaimana tiga pohon keputusan (Tree 1, Tree 2, dan Tree 3) memproses data input. Tree 1 dan Tree 3 memprediksi bahwa *instance* masuk dalam "Class A," sedangkan Tree 2 memprediksi "Class B." Dengan mekanisme voting mayoritas, hasil akhir dari Random Forest adalah "Class A," karena mendapat suara terbanyak.

<https://doi.org/10.25077/TEKNOSI.v11i2.2025.144-151>

Kinerja *Random Forest* secara teoretis bergantung pada dua faktor utama: kekuatan (*strength*, *s*) dari masing-masing pohon dan korelasi (ρ) antar pohon. Menurut Breiman [11], batas atas untuk kesalahan generalisasi (PE^*) dapat dinyatakan sebagai:

$$PE^* \leq \frac{\bar{\rho}(1 - s^2)}{s^2} \tag{1}$$

2.3. Analisis Data Eksploratif

Analisis data eksploratif dilakukan untuk memahami karakteristik data, mengidentifikasi *outlier* dengan metode *Interquartile Range* (IQR), dan menganalisis korelasi fitur untuk pemilihan fitur.

Langkah awal adalah memeriksa struktur dasar dataset. Hasil pemeriksaan pada Gambar 2 menunjukkan bahwa dataset terdiri dari 369 baris data tanpa nilai yang hilang (*null values*), sehingga tidak memerlukan proses imputasi. Statistik deskriptif pada Gambar 3 menunjukkan adanya variasi rentang yang signifikan antar fitur, misalnya Kebutuhan Kalori (rentang 849-2131) dengan Kepadatan Tulang (rentang 1.2-3.6). Perbedaan skala ini menegaskan perlunya standarisasi fitur pada tahap selanjutnya.

```

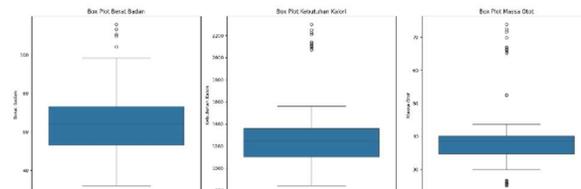
Data Info
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 369 entries, 0 to 368
Data columns (total 12 columns):
#   Column              Non-Null Count  Dtype
---  ---             
0   No                   369 non-null    int64
1   Umur                 369 non-null    int64
2   Tinggi              369 non-null    int64
3   Berat Badan         369 non-null    float64
4   Lemak Tubuh         369 non-null    float64
5   Lemak Perut         369 non-null    float64
6   Kepadatan Tulang    369 non-null    float64
7   Kebutuhan Kalori    369 non-null    int64
8   Usia Sel            369 non-null    int64
9   Massa Otot          369 non-null    float64
10  Kadar air           369 non-null    float64
11  program             369 non-null    object
dtypes: float64(6), int64(5), object(1)
memory usage: 34.7+ KB
    
```

Gambar 2. Hasil Pemeriksaan Struktur Dataset

	count	mean	std	min	max	0%	25%	50%	75%	100%
count	369	369	369	369	369	369	369	369	369	369
mean	185	29.8482	172.295	65.1414	14.5243	18.2911	2.55144	2.36287	1276.25	16.0211
std	136.681	9.53818	112.0650	14.5243	7.21945	1.11441	0.484943	0.842	141.166	17.4489
min	0	17	15	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
max	369	100	100	100	100	100	100	100	100	100
0%	0	17	15	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
25%	185	29	117	64.2	91	4.5	2.4	1.943	17	16.7
50%	227	39	169	71.8	47.2	6	3.6	1.943	63	46.1
75%	309	51	185	110.8	47.1	18	4.4	1.943	63	71.9
100%	369	100	100	100	100	100	100	100	100	100

Gambar 3. Statistik Deskriptif Variabel Numerik

Identifikasi data pencilan (*outlier*) dilakukan menggunakan metode *Interquartile Range* (IQR) yang divisualisasikan melalui *boxplot* pada Gambar 4. Hasil analisis menunjukkan adanya *outlier* pada fitur Berat Badan, Kebutuhan Kalori, dan Massa Otot.

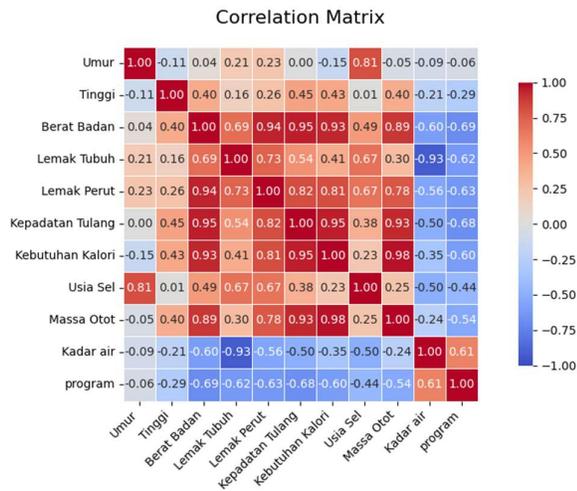


Gambar 4. Box Plot untuk Deteksi Outlier

Berdasarkan validasi dengan ahli gizi, data yang teridentifikasi sebagai *outlier* tetap dipertahankan karena merepresentasikan kasus nyata yang valid. Selain itu, sebuah percobaan dilakukan dengan melatih model pada dataset tanpa *outlier*, yang ternyata menyebabkan penurunan akurasi signifikan, yaitu dari 92,79%

menjadi 85,85%. Oleh karena itu, diputuskan untuk mempertahankan seluruh data, termasuk *outlier*, untuk melatih model yang lebih *robust* dan mampu mengenali variasi data dunia nyata [12].

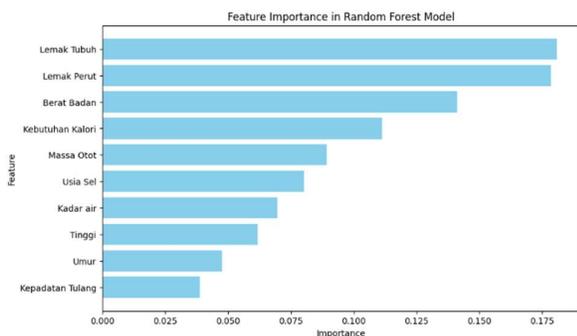
Untuk memahami hubungan antar fitur, analisis korelasi Pearson divisualisasikan menggunakan *heatmap* seperti yang ditunjukkan pada Gambar 5.



Gambar 5. Correlation Matrix

Hasilnya menunjukkan beberapa fitur memiliki hubungan linear yang cukup kuat. Contohnya, Berat Badan memiliki korelasi positif yang tinggi dengan lemak perut dan kepadatan tulang. Di sisi lain, terdapat korelasi negatif yang kuat antara lemak tubuh dan kadar air, yang mengindikasikan hubungan berbanding terbalik. Adanya korelasi kuat antar fitur ini menandakan kondisi yang disebut multikolinearitas. Kondisi ini perlu menjadi perhatian karena pada beberapa jenis model, hal ini dapat memengaruhi hasil. Namun, algoritma berbasis pohon keputusan seperti Random Forest umumnya memiliki toleransi yang lebih baik terhadap kondisi ini.

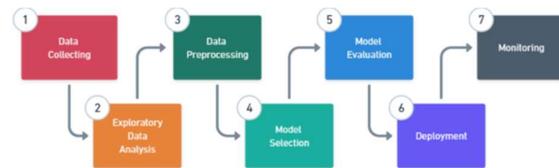
Selanjutnya, analisis kepentingan fitur (*feature importance*) dilakukan menggunakan algoritma *Random Forest* seperti yang ditunjukkan pada Gambar 6. Hasilnya menunjukkan bahwa semua fitur memiliki nilai kepentingan di atas nol, dengan Lemak Tubuh, Lemak Perut, dan Berat Badan sebagai fitur paling berpengaruh.



Gambar 6. Feature Importance

Berdasarkan masukan ahli gizi bahwa semua fitur berguna untuk menentukan program diet, serta temuan bahwa semua fitur memiliki kontribusi, diputuskan untuk menggunakan seluruh 11 fitur prediktor dalam pembangunan model. Keputusan ini didukung oleh literatur yang menyatakan bahwa fitur yang berkorelasi tinggi dengan fitur penting dapat dianggap relevan sebagai bagian dari kolaborasi fitur yang saling mendukung [13]. Hal ini mengindikasikan bahwa meskipun fitur tersebut tidak terlalu dominan secara langsung, mereka tetap berperan dalam sinergi fitur lain yang lebih penting.

2.4. Pra-pemrosesan Data



Gambar 7. Diagram Blok Proses Pengolahan Data

Tahap pra-pemrosesan data merupakan langkah krusial untuk membersihkan dan mempersiapkan dataset agar siap digunakan dalam pelatihan model *machine learning*. Proses ini memastikan bahwa data yang digunakan berkualitas tinggi, konsisten, dan memiliki format yang sesuai, sehingga dapat meningkatkan akurasi dan keandalan model. Tahapan yang dilakukan meliputi *encoding* label, pembagian data secara terstratifikasi, dan normalisasi fitur.

2.4.1. Encoding Label Target

Langkah pertama adalah mengubah variabel target program yang bersifat kategorikal (teks) menjadi representasi numerik. Transformasi ini merupakan prasyarat wajib agar algoritma *machine learning* seperti *Random Forest* dapat memproses dan mempelajari data target. Dalam penelitian ini, pemetaan (*mapping*) dilakukan secara manual sebagai berikut:

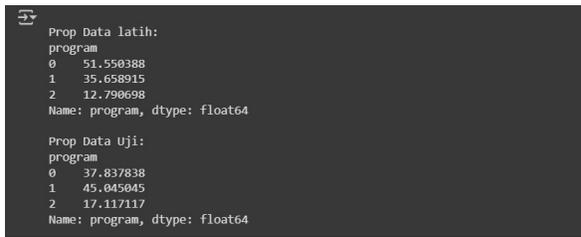
1. 'Turun BB' diubah menjadi 0
2. 'Turun Lemak' diubah menjadi 1
3. 'Naik BB' diubah menjadi 2

Hasil dari *encoding* ini adalah kolom target yang sepenuhnya numerik dan siap untuk digunakan dalam proses selanjutnya.

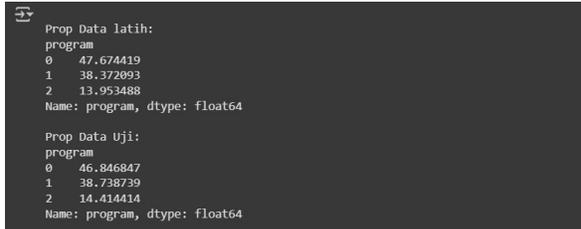
2.4.2. Pembagian Dataset

Dataset selanjutnya dibagi menjadi data latih dan data uji dengan proporsi 70:30. Dari total 369 sampel, sebanyak 258 data digunakan untuk melatih model (data latih) dan 111 data digunakan untuk menguji performa model (data uji).

Proses pembagian ini dilakukan menggunakan fungsi *train_test_split* dari pustaka *Scikit-learn* dengan parameter *stratify=y*. Parameter tersebut digunakan untuk memastikan bahwa distribusi kelas target ('Turun BB', 'Turun Lemak', 'Naik BB') tetap proporsional antara data latih dan data uji. Pendekatan *stratified random sampling* ini sangat penting untuk mencegah bias, terutama pada dataset yang memiliki ketidakseimbangan kelas, sehingga model dapat dilatih dan dievaluasi pada data yang representatif. Hasil dari penerapan stratifikasi ini sangat signifikan, seperti yang ditunjukkan pada perbandingan Gambar 8 dan Gambar 9 di bawah ini.



Gambar 8. Distribusi Label tanpa Stratifikasi



Gambar 9. Distribusi Label dengan Stratifikasi

Pada Gambar 8 distribusi kelas menjadi tidak seimbang tanpa adanya stratifikasi. Misalnya proporsi kelas 0 (Turun Berat Badan) pada data latih sebesar 51.55%, sangat berbeda dengan data uji yaitu 37.83%. Sebaliknya, dengan stratifikasi seperti yang ditunjukkan pada Gambar 9, distribusi kelas menjadi jauh lebih seimbang dan representatif, di mana proporsi kelas 0 pada data latih sebesar 47.67%, hampir sama dengan pada data uji sebesar 46.85%.

2.4.3. Standarisasi Fitur (Feature Standarization)

Analisis statistik deskriptif pada data mentah menunjukkan adanya rentang nilai yang sangat beragam antar fitur. Misalnya, fitur Kebutuhan Kalori memiliki nilai dalam ribuan (rata-rata 1243.6), sedangkan Kepadatan Tulang memiliki nilai dalam satuan (rata-rata 2.3). Perbedaan skala yang signifikan ini dapat menyebabkan model memberikan bobot yang lebih besar pada fitur dengan rentang nilai yang lebih tinggi, sehingga berpotensi mengabaikan kontribusi fitur dengan nilai yang lebih kecil.

Untuk mengatasi masalah ini, dilakukan standarisasi pada seluruh fitur prediktor menggunakan StandardScaler dari pustaka 1.

Scikit-learn. Metode ini mentransformasi setiap fitur sehingga memiliki rata-rata (mean) 0 dan standar deviasi 1. Dengan demikian, semua fitur memiliki skala yang seragam dan sebanding, memungkinkan model untuk mengevaluasi kontribusi setiap fitur secara adil selama proses pelatihan. Hasil dari tahap ini adalah dataset fitur (X_{train_scaled} dan X_{test_scaled}) yang siap digunakan untuk melatih model secara optimal.

2.5. Pembangunan Model

Model prediksi dibangun menggunakan algoritma Random Forest Classifier. Pemilihan ini didasarkan pada kemampuannya menangani interaksi kompleks antar fitur dan ketahanannya terhadap *overfitting*. Untuk optimasi performa, dilakukan pencarian parameter terbaik (*hyperparameter tuning*) menggunakan GridSearchCV dengan validasi silang *10-fold*. Sehingga, Parameter terbaik yang diperoleh adalah $n_estimators=500$, $max_depth=10$, $min_samples_split=2$, $min_samples_leaf=1$, dengan $class_weight='balanced'$.

Kinerja model dievaluasi menggunakan metrik standar untuk masalah klasifikasi, yaitu akurasi, presisi, *recall*, dan *F1-score*. Rumus yang digunakan adalah sebagai berikut:

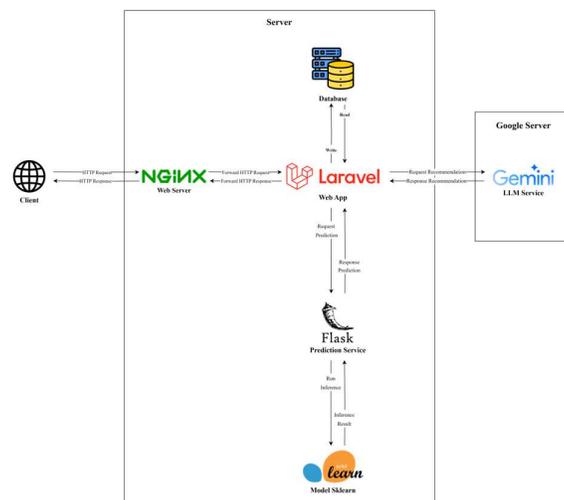
$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Sampel} = \frac{\sum TP}{\sum TP + \sum FP + \sum FN + \sum TN} \quad (2)$$

$$Presisi_i = \frac{TP_i}{TP_i + FP_i} \quad (3)$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (4)$$

$$F - 1\ Score_i = \frac{2 \times Precision_i \times Recall_i}{Precision_i + Recall_i} \quad (5)$$

Di mana TP (*True Positive*), TN (*True Negative*), FP (*False Positive*), dan FN (*False Negative*) adalah elemen dari *confusion matrix*.



Gambar 10. Diagram Arsitektur Sistem

2.6. Integrasi Large Language Models (LLM)

Integrasi dengan LLM bertujuan untuk menghasilkan rekomendasi diet naratif yang dinamis dan personal. Proses ini diawali dengan penyusunan prompt secara otomatis oleh sistem setelah mendapatkan hasil prediksi program diet. Prompt ini menggabungkan beberapa elemen kunci:

2. Konteks peran yang merupakan instruksi awal yang mendefinisikan peran LLM sebagai seorang ahli gizi yang akan memberikan rekomendasi diet kepada pelanggan.
3. Data pelanggan berupa data komposisi tubuh pelanggan sesuai hasil *checkup* terkini.
4. Hasil prediksi program diet yang dipresiksi oleh model Random Forest.
5. Preferensi pengguna berupa informasi tambahan seperti alergi atau pilihan makanan halal atau non-halal.
6. Template output berupa struktur keluaran yang diinginkan, mencakup saran makanan, teknik pengolahan makanan, anjuran nutrisi, dan aktivitas fisik.

Arsitektur sistem secara umum dapat dilihat Gambar 10. Permulaan *Prompt* yang telah tersusun kemudian dikirimkan ke *endpoint* Gemini Flash-Lite API melalui permintaan HTTP POST. Parameter temperature diatur pada nilai 0.7 untuk menghasilkan keluaran yang kreatif namun tetap konsisten, dan *maxOutputTokens* sebesar 2048 untuk memastikan rekomendasi yang dihasilkan cukup detail. Respons dari API yang berisi teks rekomendasi kemudian disimpan di database untuk ditampilkan kepada pengguna.

2.7. Metode Pengujian

Pengujian sistem dilakukan secara multi-aspek dan komprehensif untuk mengukur keandalan dari berbagai sisi, mulai dari performa teknis model hingga penerimaan oleh pengguna akhir.

2.7.1. Pengujian Kinerja Model Random Forest

Pengujian ini bertujuan untuk mengukur akurasi dan kemampuan generalisasi. Pengujian dilakukan dalam dua tahap: pertama, evaluasi pada 30% data uji (111 sampel) menggunakan metrik standar klasifikasi (akurasi, presisi, *recall*, dan *F1-score*) yang dihitung dari *confusion matrix*. Kedua, untuk menguji performa pada data dunia nyata yang sepenuhnya baru (*unseen data*), dilakukan pengujian pada 50 data pelanggan asli dari mitra. Hasil prediksi sistem pada tahap ini dibandingkan dengan label acuan (*ground truth*) yang ditentukan oleh asisten ahli gizi.

2.7.2. Pengujian Fungsionalitas

Pengujian fungsionalitas sistem menggunakan metode *black box testing* untuk memverifikasi bahwa setiap fitur aplikasi berfungsi sesuai spesifikasi, tanpa memperhatikan struktur kode internal. Pengujian dilakukan dengan menjalankan skenario-skenario penggunaan pada 10 modul utama sistem, seperti otentikasi pengguna, manajemen data (CRUD), alur prediksi, integrasi LLM, dan pembuatan laporan. Sebuah fitur dinyatakan berhasil jika output yang dihasilkan sesuai dengan output yang diharapkan untuk setiap input yang diberikan.

2.7.3. Pengujian System Usability Scale (SUS)

Tingkat kemudahan penggunaan dan kepuasan pengguna diukur menggunakan metode *System Usability Scale* (SUS). Kuesioner standar SUS yang terdiri dari 10 pernyataan dengan skala Likert 5 poin disebarakan kepada 15 responden yang mencakup seluruh peran pengguna: 1 ahli gizi, 1 asisten ahli gizi, dan 13 pelanggan. Skor akhir dihitung dengan menjumlahkan kontribusi dari setiap pertanyaan dan dikalikan 2.5 untuk menghasilkan nilai dalam rentang 0-100, yang kemudian diinterpretasikan menggunakan skala standar SUS.

Metode ini menggunakan kuesioner standar yang berisi 10 pernyataan yang disajikan pada _ untuk mengevaluasi persepsi pengguna secara subjektif.

Tabel 1. Pernyataan dalam Kuesioner SUS

No	Pernyataan
1	Saya merasa sistem ini mudah digunakan.
2	Saya merasa sistem ini terlalu kompleks.
3	Saya merasa sistem ini mudah dipelajari.
4	Saya membutuhkan bantuan teknis untuk bisa menggunakan sistem ini.
5	Saya merasa fitur dalam sistem ini sudah terintegrasi dengan baik
6	Saya merasa ada terlalu banyak inkonsistensi dalam sistem ini.
7	Saya yakin bahwa sebagian besar orang bisa belajar menggunakan sistem ini dengan cepat.
8	Saya merasa sistem ini membingungkan ketika digunakan.
9	Saya merasa nyaman menggunakan sistem ini.
10	Saya membutuhkan banyak pelatihan sebelum bisa menggunakan sistem ini.

2.7.4. Pengujian Rekomendasi LLM

Validasi ini secara spesifik mengevaluasi kualitas, relevansi, dan kelengkapan rekomendasi diet yang dihasilkan oleh LLM (Gemini Flash-Lite). Lima sampel rekomendasi yang dihasilkan sistem dari data pelanggan baru divalidasi oleh ahli gizi menggunakan kuesioner terstruktur. Penilaian berfokus pada tiga aspek utama: (1) kesesuaian dengan kebutuhan pelanggan, (2) kelengkapan cakupan aspek diet (pola makan, porsi, aktivitas fisik), dan (3) kelayakan sebagai referensi awal. Persentase kelayakan dihitung untuk memberikan ukuran kuantitatif dari kualitas output LLM.

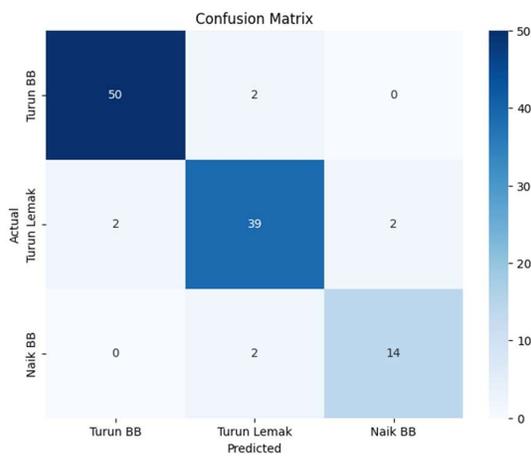
3. HASIL

Pada penelitian ini, evaluasi mencakup dua area utama: pertama, analisis kuantitatif terhadap kinerja model *machine learning* (Random Forest) dalam memprediksi program diet, Kedua, pengujian sistem aplikasi secara menyeluruh, yang meliputi verifikasi fungsionalitas dengan metode *black box*, pengukuran tingkat kemudahan penggunaan (*usability*) dari perspektif pengguna melalui *System Usability Scale* (SUS), serta validasi kualitas rekomendasi diet yang dihasilkan oleh integrasi *Large Language Model* (LLM).

3.1. Kinerja Model Prediksi

Model *Random Forest* yang telah dioptimalkan menunjukkan kinerja yang sangat baik pada data uji. Setelah proses pelatihan selesai, model dievaluasi terhadap data uji menggunakan *confusion matrix* serta metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Akurasi keseluruhan yang dicapai adalah 93%.

Dari hasil tersebut menunjukkan bahwa model mampu memprediksi program diet dengan tingkat kesalahan yang rendah. Nilai-nilai yang tinggi pada setiap metrik ini menandakan bahwa model tidak hanya akurat dalam mengklasifikasi setiap label, tetapi juga seimbang dalam mendeteksi masing-masing kategori program diet tanpa kecenderungan bias terhadap kelas tertentu. Analisis lebih dalam menggunakan *confusion matrix* yang ditunjukkan pada Gambar 11 dan laporan klasifikasi pada Tabel 2.



Gambar 11. *Confusion Matrix*

Tabel 2. Laporan Klasifikasi Performa Model

Label	Precision	Recall	F1-score	Jumlah Data
0 (Turun BB)	0.96	0.96	0.96	52
1 (Turun Lemak)	0.91	0.91	0.91	43
2 (Naik BB)	0.88	0.88	0.88	16
Akurasi			0.93	111

Selain itu, pengujian pada 50 data baru dari mitra yang tidak termasuk dalam dataset pelatihan menghasilkan akurasi sebesar 84%. Hasil validasi silang *10-fold* juga menunjukkan performa yang konsisten dengan rata-rata akurasi 91%, menandakan model tidak mengalami *overfitting* dan memiliki kemampuan generalisasi yang baik.

3.2. Pengujian Sistem

Pengujian sistem dilakukan untuk mengevaluasi aspek-aspek di luar performa inti model, yaitu fungsionalitas aplikasi, kemudahan penggunaan (*usability*), dan kualitas rekomendasi yang dihasilkan.

Pengujian fungsionalitas dilakukan menggunakan metode *black box* dan menunjukkan bahwa 100% skenario pengujian berhasil. Seluruh fitur utama sistem, termasuk otentikasi, manajemen data, alur prediksi, dan integrasi API, berjalan sesuai dengan spesifikasi.

Hasil pengujian *usability* terhadap 15 responden menghasilkan skor rata-rata SUS sebesar 85,17. Berdasarkan skala interpretasi standar Tabel 3, skor ini termasuk dalam kategori Sangat Baik.

Tabel 3. Interpretasi Skor SUS

Skor SUS	Keterangan
81 – 100	Sangat Baik
69 - 80	Baik
51 - 68	Marginal
0 - 50	Tidak Dapat Diterima

Tabel 4. Hasil Validasi Kualitas Rekomendasi LLM

Aspek Penilaian	Skor (1-5)
Kesesuaian rekomendasi dengan kebutuhan pelanggan	4
Kelengkapan aspek diet yang dicakup	5
Kelayakan sebagai referensi awal	4
Persentase Kelayakan	86,6%
Keterangan	Sangat Layak

Hal ini menunjukkan bahwa sistem memiliki tingkat kemudahan penggunaan yang tinggi dan secara umum dapat diterima oleh pengguna secara umum.

Validasi oleh ahli gizi terhadap 5 sampel rekomendasi diet menghasilkan nilai kelayakan keseluruhan sebesar 86,6%, yang

dikategorikan sebagai Sangat Layak, seperti yang disajikan dalam Tabel 4.

4. PEMBAHASAN

Hasil penelitian ini menunjukkan bahwa implementasi sistem terintegrasi berhasil memenuhi tujuan yang ditetapkan. Model

prediksi *Random Forest* dengan akurasi 93% pada data uji dan 84% pada data baru mengonfirmasi efektivitasnya dalam domain prediksi kesehatan, sejalan dengan temuan pada penelitian sebelumnya [5] dan [14] yang menyatakan bahwa *Random Forest* termasuk salah satu model dengan akurasi tertinggi untuk klasifikasi obesitas, dengan performa lebih baik dibandingkan berbagai algoritma pembelajaran mesin lainnya yang diuji dalam penelitian mereka. Tingkat akurasi ini membuktikan kemampuan model untuk mengenali pola kompleks dari data komposisi tubuh, menjadikannya alat bantu yang andal bagi ahli gizi.

Inovasi utama penelitian ini terletak pada integrasi hasil prediksi ML dengan LLM untuk menghasilkan rekomendasi diet yang dipersonalisasi. Tingkat kelayakan rekomendasi sebesar 86,6% yang divalidasi oleh ahli gizi menunjukkan bahwa LLM mampu menerjemahkan data kuantitatif menjadi saran praktis yang relevan dan komprehensif. Hal ini mendukung temuan Szymanski dkk. [14] mengenai potensi LLM sebagai asisten nutrisi. Meskipun demikian, seperti yang dicatat oleh Hamarashid dkk. [12], peran validasi oleh ahli gizi tetap krusial untuk memastikan akurasi klinis dan menghindari potensi "halusinasi" dari LLM.

Dari sisi pengguna, skor *usability* SUS sebesar 85,17 (Sangat Baik) menunjukkan bahwa sistem berhasil dirancang dengan antarmuka yang efektif dan mudah diterima. Sistem yang dikembangkan secara efektif menjembatani kesenjangan penelitian sebelumnya dengan menyediakan solusi *end-to-end*—mulai dari digitalisasi data, prediksi, hingga rekomendasi.

5. KESIMPULAN

Berdasarkan hasil penelitian dan pengujian, dapat ditarik tiga kesimpulan utama. Pertama, sistem pencatatan data digital terbukti efektif mengurangi risiko kehilangan data dan meningkatkan efisiensi analisis data historis dibandingkan metode manual. Kedua, model prediksi berbasis *Random Forest* berhasil mengklasifikasikan program diet dengan akurasi tinggi, yaitu sebesar 93% pada data uji, dan kemampuan generalisasi yang baik sebesar 84% pada data baru. Ketiga, integrasi dengan LLM mampu menghasilkan rekomendasi diet yang relevan dan komprehensif, divalidasi dengan tingkat kelayakan 86,6%, dan didukung oleh tingkat *usability* sistem yang sangat baik dengan skor SUS 85,17.

Untuk pengembangan selanjutnya, disarankan untuk memperluas platform ke dalam bentuk aplikasi *mobile* guna meningkatkan aksesibilitas dan kemudahan interaksi bagi pelanggan. Selain itu, perlu dilakukan eksplorasi dan perbandingan dengan algoritma *machine learning* lain untuk memvalidasi apakah terdapat model alternatif yang dapat memberikan akurasi atau efisiensi komputasi yang lebih tinggi. Dari sisi rekomendasi, pengembangan mekanisme umpan balik (*feedback loop*) yang memungkinkan sistem belajar dari koreksi ahli gizi, serta penambahan *knowledge base* khusus, dapat secara signifikan meningkatkan akurasi dan personalisasi rekomendasi yang dihasilkan oleh LLM.

DAFTAR PUSTAKA

- [1] World Health Organization, "Malnutrition," 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/malnutrition?gad_source=1&gclid=CjwKCAiA6aW6BhBqEiwA6KzDcyRGe2D05CSMaJqYt-OMYtVZtAOrE8e04gwzGFUEt50hysgvK6bRERoCyx oQAvD_BwE#
- [2] R. Rosmiati, N. R. Haryana, H. Firmansyah, and R. Purba, "Pola Makan, Aktivitas Fisik, dan Obesitas pada Pekerja Urban di Indonesia," *Amerta Nutrition*, vol. 7, pp. 164–170, Dec. 2023, doi: [10.20473/amnt.v7i2SP.2023.164-170](https://doi.org/10.20473/amnt.v7i2SP.2023.164-170).
- [3] J. S. Shim, S. Y. Shim, H. J. Cha, J. Kim, and H. C. Kim, "Association between Ultra-processed Food Consumption and Dietary Intake and Diet Quality in Korean Adults," *J Acad Nutr Diet*, vol. 122, no. 3, pp. 583–594, Mar. 2022, doi: [10.1016/j.jand.2021.07.012](https://doi.org/10.1016/j.jand.2021.07.012).
- [4] A. Setyaningsih, I. Mulyasari, P. Afiatna, and H. R. Putri, "The Relationship between Ultra-Processed Food Consumption with Diet Quality and Overweight Status in Young Adults," *Amerta Nutrition*, vol. 8, no. 1, pp. 124–129, 2024, doi: [10.20473/amnt.v8i1.2024.124-129](https://doi.org/10.20473/amnt.v8i1.2024.124-129).
- [5] N. C. Pereira, J. D'Souza, P. Rana, and S. Solaskar, "Obesity Related Disease Prediction from Healthcare Communities Using Machine Learning," *2019 10th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2019*, pp. 1–7, 2019, doi: [10.1109/ICCCNT45670.2019.8944798](https://doi.org/10.1109/ICCCNT45670.2019.8944798).
- [6] E. Rodríguez, E. Rodríguez, L. Nascimento, A. da Silva, and F. Marins, "Machine learning techniques to predict overweight or obesity," *CEUR Workshop Proc*, vol. 3038, pp. 190–204, 2021.
- [7] M. Dirik, "Application of machine learning techniques for obesity prediction: a comparative study," *Journal of Complexity in Health Sciences*, vol. 6, no. 2, pp. 16–34, 2023, doi: [10.21595/chs.2023.23193](https://doi.org/10.21595/chs.2023.23193).
- [8] V. Sharma *et al.*, "Malnutrition, Health and the Role of Machine Learning in Clinical Setting," *Front Nutr*, vol. 7, no. April, pp. 1–9, 2020, doi: [10.3389/fnut.2020.00044](https://doi.org/10.3389/fnut.2020.00044).
- [9] A. S. Iorhemen, "Random Forest Ensemble Machine Learning Model for Early Detection and Prediction of Weight Category," *Journal of Data Science and Intelligent Systems*, vol. 2, no. May 2023, pp. 233–240, 2023, doi: [10.47852/bonviewjdsis32021149](https://doi.org/10.47852/bonviewjdsis32021149).
- [10] C. Suresh, B. V. Kiranmayee, M. Jahnavi, R. Pampari, S. R. Ambadipudi, and S. S. P. Hemadri, "Obesity Prediction Based on Daily Lifestyle Habits and Other Factors Using Different Machine Learning Algorithms," *Proceedings of Second International Conference on Advances in Computer Engineering and Communication Systems*, pp. 397–407, 2022, doi: [10.1007/978-981-16-7389-4_39](https://doi.org/10.1007/978-981-16-7389-4_39).
- [11] L. Breiman, "Random Forest," *Kluwer Academic Publishers*, vol. 45, pp. 5–32, 2001.
- [12] J. Yang *et al.*, "Mitigating machine learning bias between high income and low–middle income countries for enhanced model fairness and generalizability," *Sci* <https://doi.org/10.25077/TEKNOSI.v11i2.2025.144-151>

- Rep*, vol. 14, no. 1, pp. 1–12, 2024, doi: [10.1038/s41598-024-64210-5](https://doi.org/10.1038/s41598-024-64210-5).
- [13] Y. Rothacher and C. Strobl, *Identifying Informative Predictor Variables With Random Forests*, vol. 49, no. 4, 2024. doi: [10.3102/10769986231193327](https://doi.org/10.3102/10769986231193327).
- [14] A. Szymanski, B. L. Wimer, O. Anuyah, H. A. Eicher-Miller, and R. A. Metoyer, “Integrating Expertise in LLMs: Crafting a Customized Nutrition Assistant with Refined Template Instructions,” *Conference on Human Factors in Computing Systems - Proceedings*, 2024, doi: [10.1145/3613904.3641924](https://doi.org/10.1145/3613904.3641924).