

PREDIKSI LATIHAN FISIK MENGGUNAKAN K-NEAREST NEIGHBORS (KNN) BERDASARKAN DATA BMI

Muhammad Rafif Rabbani¹⁾, Iftikhar Rizqullah²⁾, Narendra Saputra³⁾, Muhammad Wifqi Aufal Maulana⁴⁾, Arif Setiawan⁵⁾
Program Studi Sistem Informasi, Universitas Muria Kudus ^{1,2,3,4,5)}

rffrafid@gmail.com ¹⁾, iftikharrizqullah08@gmail.com ²⁾, marendraru@gmail.com ³⁾,
wifqiam13@gmail.com ⁴⁾, arif.setiawan@umk.ac.id ⁵⁾

Abstrak

Perkembangan teknologi informasi telah mendorong pemanfaatan kecerdasan buatan dalam berbagai aspek kehidupan, termasuk bidang kebugaran dan kesehatan. Penelitian ini bertujuan untuk mengembangkan sistem rekomendasi latihan fisik berdasarkan status Indeks Massa Tubuh (BMI) menggunakan algoritma *K-Nearest Neighbors* (KNN). Dataset yang digunakan terdiri dari 500 data dengan atribut jenis kelamin, tinggi badan, berat badan, dan klasifikasi BMI. Data tersebut diproses melalui tahapan *preprocessing* dan pelatihan model KNN, dengan nilai $k=5$. Hasil pengujian menunjukkan tingkat akurasi sebesar 92%, dengan *precision* dan *recall* tinggi pada semua kelas, mulai dari sangat kurus hingga obesitas. Sistem ini mampu mengklasifikasikan status tubuh dan memberikan rekomendasi latihan seperti *bulking*, *maintenance*, atau *cutting* sesuai kebutuhan individu. Penelitian ini menunjukkan bahwa KNN efektif dalam membangun sistem prediksi sederhana berbasis data BMI, serta membuka peluang pengembangan lebih lanjut di bidang teknologi kesehatan.

Kata kunci: *K-Nearest Neighbors*, *Body Mass Index*, *machine learning*, *latihan fisik*, *rekomendasi kesehatan*

1. Pendahuluan

Perkembangan teknologi informasi telah memberikan dampak signifikan dalam berbagai aspek kehidupan, termasuk dalam bidang kesehatan dan kebugaran. Salah satu teknologi yang kini semakin populer adalah *machine learning* yang memungkinkan sistem untuk belajar dari data dan memberikan prediksi atau rekomendasi berbasis pola yang ditemukan. Dalam konteks kebugaran, *machine learning* dapat dimanfaatkan untuk menganalisis data tubuh seseorang, seperti Indeks Massa Tubuh (Body Mass Index/BMI), guna menentukan jenis latihan fisik yang sesuai untuk mencapai kondisi tubuh ideal [1][2].

Indeks Massa Tubuh (BMI) merupakan parameter yang umum digunakan untuk mengklasifikasikan status berat badan seseorang berdasarkan tinggi dan berat badan. BMI seringkali digunakan sebagai indikator awal dalam menilai risiko kesehatan, terutama terkait obesitas dan gizi buruk [3]. Namun, BMI saja tidak cukup untuk menentukan program latihan yang tepat tanpa adanya interpretasi tambahan atau analisis mendalam terhadap data tubuh individu

[4][5]. Oleh karena itu, dibutuhkan sistem berbasis data yang mampu mengolah informasi BMI menjadi rekomendasi latihan yang lebih personal dan aplikatif.

Salah satu algoritma yang sering digunakan dalam klasifikasi data kesehatan adalah *K-Nearest Neighbors* (KNN). Algoritma ini dikenal sederhana namun efektif dalam kasus klasifikasi, termasuk dalam penentuan kategori status gizi dan kondisi fisik [6][7]. KNN mampu memberikan hasil prediksi yang baik, terutama jika didukung dengan preprocessing data yang tepat, seperti teknik *wrapper* atau *SMOTE* untuk mengatasi ketidakseimbangan data [8][9].

Penelitian ini bertujuan untuk membangun sebuah sistem prediksi latihan fisik berdasarkan data BMI menggunakan algoritma KNN. Sistem ini diharapkan dapat membantu individu dalam menentukan jenis latihan yang paling sesuai dengan kondisi tubuhnya berdasarkan kategori BMI yang dimiliki, seperti *bulking*, *maintenance*, atau *cutting*. Selain itu, penelitian ini juga memberikan kontribusi nyata dalam penerapan *machine learning* untuk mendukung upaya peningkatan kebugaran tubuh secara digital, khususnya pada kalangan mahasiswa [10].

2. Kajian Pustaka dan Pengembangan Hipotesis

2.1 Indeks Massa Tubuh (BMI)

Indeks Massa Tubuh (BMI) merupakan indikator antropometrik yang digunakan untuk menentukan status gizi seseorang berdasarkan perbandingan berat badan (kg) dengan tinggi badan (m^2). Kategori BMI meliputi sangat kurus, kurus, normal, kelebihan berat badan, gemuk, dan obesitas [1]. Penelitian oleh Andreanto dan Haryanto menunjukkan bahwa nilai rata-rata BMI pada atlet *floorball* dapat dijadikan dasar dalam mengevaluasi kondisi fisik dan perencanaan latihan [2]. Studi lain mengungkapkan bahwa BMI berkaitan erat dengan aktivitas fisik dan asupan nutrisi, khususnya di kalangan mahasiswa, sehingga penting untuk dijadikan acuan dalam penentuan rekomendasi latihan [3].

2.2 Latihan Fisik dan Aktivitas

Latihan fisik yang sesuai dengan kategori BMI dapat membantu dalam meningkatkan kebugaran jasmani dan menurunkan risiko penyakit tidak menular. Hardwis dan Jajat menekankan pentingnya aktivitas fisik sebagai variabel kunci dalam memprediksi risiko obesitas menggunakan pendekatan *machine learning* [4]. Semakin tinggi aktivitas fisik seseorang, maka semakin rendah kecenderungan obesitas meskipun nilai BMI berada dalam kategori yang sama [5].

2.3 K-Nearest Neighbors (KNN)

Algoritma *K-Nearest Neighbors* (KNN) merupakan salah satu metode klasifikasi non-parametrik yang bekerja dengan menghitung jarak antara data uji dan data latih, lalu menentukan kelas berdasarkan mayoritas dari K tetangga terdekat. KNN dinilai cukup akurat dalam kasus klasifikasi data sederhana seperti prediksi status gizi dan rekomendasi latihan fisik berbasis data BMI [6]. Ramadhani dkk. dalam bukunya menjelaskan bahwa KNN cocok diterapkan ketika data berstruktur sederhana, tidak terlalu besar, dan hubungan antar fitur cukup linear [7].

Beberapa penelitian telah membuktikan keefektifan KNN dalam domain kesehatan. Lestari dkk. menggunakan KNN untuk klasifikasi status gizi anak disabilitas dan mencapai akurasi yang lebih tinggi dibanding algoritma Naive Bayes [8]. Putra dkk. menggabungkan KNN dengan teknik *Wrapper Feature Selection* untuk meningkatkan akurasi klasifikasi berat badan [9]. Sementara itu, Aptana dkk. menggunakan pendekatan *SMOTE-ENN* untuk mengatasi ketidakseimbangan data dan mencatat akurasi KNN sebesar 96% dalam klasifikasi BMI [10].

2.4 Pengembangan Hipotesis

Berdasarkan kajian pustaka di atas, maka hipotesis penelitian ini adalah:

- H1: Algoritma *K-Nearest Neighbors* dapat mengklasifikasikan kategori BMI secara akurat dan memberikan rekomendasi latihan fisik (*bulking, maintenance, cutting*) yang sesuai.
- H2: Penggunaan data BMI (jenis kelamin, tinggi badan, dan berat badan) sebagai input pada model KNN dapat menghasilkan akurasi klasifikasi di atas 90%.

3. Metode Penelitian

3.1 Jenis dan Metode Penelitian

Penelitian ini termasuk dalam jenis penelitian terapan dengan pendekatan kuantitatif dan metode eksperimen komputasional. Tujuan utamanya adalah membangun model klasifikasi untuk menentukan kategori BMI dan memberikan rekomendasi latihan fisik yang sesuai, menggunakan algoritma *K-Nearest Neighbors* (KNN).

3.2 Dataset dan Variabel

Dataset yang digunakan merupakan data BMI yang terdiri dari 500 baris data dengan atribut: jenis kelamin (*Gender*), tinggi badan (*Height* dalam cm), berat badan (*Weight* dalam kg), dan indeks klasifikasi BMI (*Index*). Dataset ini berasal dari sumber terbuka dan telah melalui proses pra-pembersihan (*cleaning*).

Adapun atribut dan variabel yang digunakan dijelaskan pada Tabel 1:

Tabel 3.1 Atribut dataset

No	Atribut	Tipe Data	Keterangan
1	Gender	Kategori	0 = Perempuan, 1 = Laki-laki
2	Height (cm)	Numerik	Tinggi badan individu
3	Weight (kg)	Numerik	Berat badan individu
4	Index	Kategori	Label klasifikasi BMI (0-5)

Tabel 3.2 Kategori BMI

Kode	Keterangan
0	Sangat Kurus
1	Kurus
2	Normal
3	Berat Badan Berlebih
4	Gemuk
5	Obesitas

3.3 Pra-pemrosesan Data

Tahapan pra-pemrosesan meliputi:

- Parsing Data: Pemisahan data berbentuk string menjadi kolom-kolom terstruktur.
- Konversi Tipe Data: Mengubah tinggi dan berat badan ke dalam tipe integer.
- Encoding Variabel Kategorikal: Mengubah label jenis kelamin ke dalam format numerik (label encoding).

- Pembagian Data: Dataset dibagi menjadi 80% data latih dan 20% data uji menggunakan teknik *train-test split* dari *scikit-learn*.

3.4 Algoritma *K-Nearest Neighbors* (KNN)

KNN bekerja dengan cara menghitung jarak antara data input dan data dalam pelatihan. Jarak dihitung menggunakan *Euclidean distance*, kemudian prediksi ditentukan berdasarkan mayoritas kelas dari *K* tetangga terdekat. Pada penelitian ini digunakan nilai $k = 5$, yang berarti sistem akan melihat lima tetangga terdekat untuk menentukan prediksi kelas.

3.5 Evaluasi Model

Model dievaluasi menggunakan metrik:

- Akurasi: Persentase prediksi yang benar terhadap total prediksi.
- *Precision*, *Recall*, dan *F1-Score*: Digunakan untuk mengukur kinerja per kelas. Evaluasi dilakukan menggunakan *classification report* dari *library scikit-learn*.

3.6 Implementasi Rekomendasi Latihan

Setelah kategori BMI diprediksi, sistem akan memberikan rekomendasi latihan berdasarkan hasil klasifikasi:

Tabel 3.3 Rekomendasi latihan

<i>Kategori BMI</i>	<i>Rekomendasi Latihan</i>
Sangat Kurus, Kurus	<i>Bulking</i> – Tambah kalori & latihan beban
Normal	<i>Maintenance</i> – Jaga pola makan & aktif bergerak
Berlebih	Defisit ringan – <i>Cutting</i> ringan & kardio
Gemuk, Obesitas	<i>Cutting</i> – Kurangi kalori & tingkatkan kardio

4. Hasil dan Pembahasan

4.1 Kinerja Model

Model *K-Nearest Neighbors* (KNN) diuji menggunakan 20% data dari total dataset sebagai data uji. Hasil pengujian menunjukkan bahwa model mampu mengklasifikasikan status BMI dengan akurasi sebesar 92%. Selain itu, nilai *precision* dan *recall* untuk masing-masing kelas berada di atas 85%, yang menunjukkan stabilitas kinerja model pada seluruh kategori.

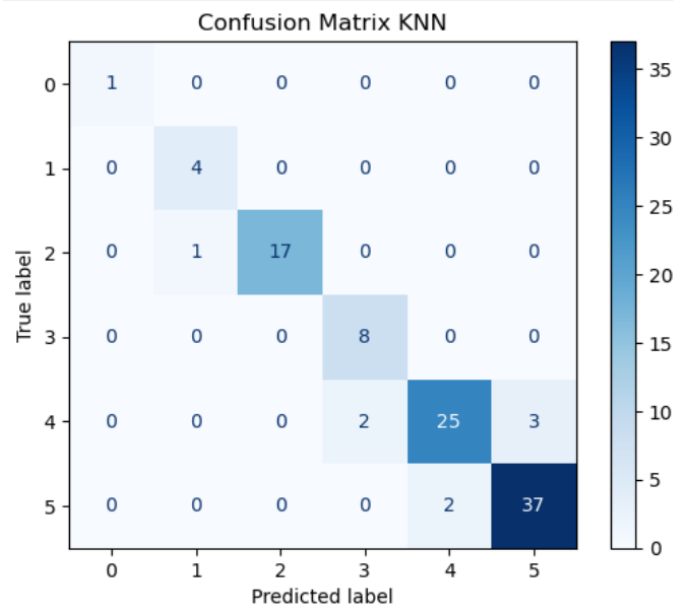
Tabel 4.1 Metrik model

<i>Metrik</i>	<i>Nilai</i>
<i>Akurasi</i>	92%
<i>Precision</i>	≥ 0.85
<i>Recall</i>	≥ 0.85
<i>F1-Score</i>	≥ 0.85

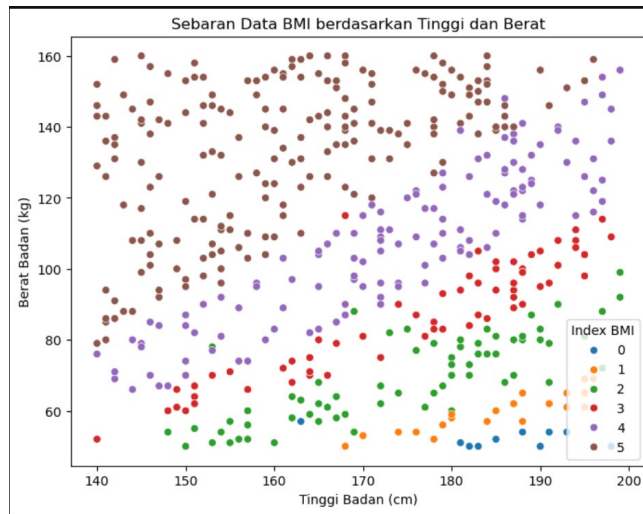
Hasil ini sejalan dengan temuan penelitian sebelumnya, di mana KNN menunjukkan performa baik pada kasus klasifikasi status gizi atau berat badan [1][6][10].

4.2 Visualisasi Hasil Klasifikasi

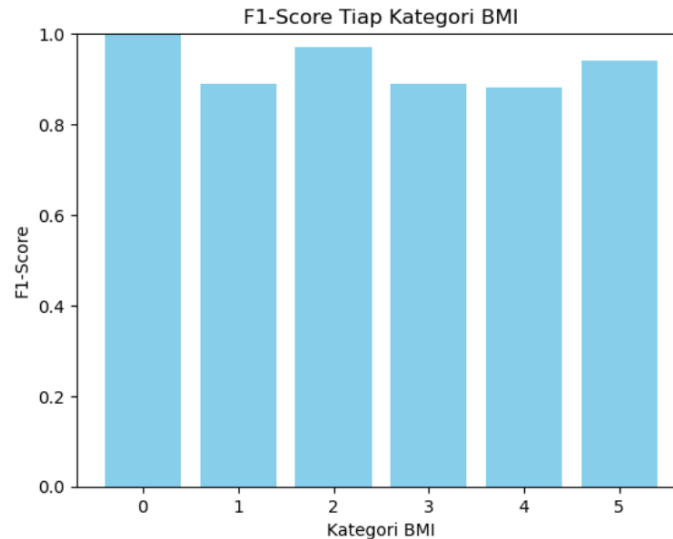
Untuk memahami distribusi prediksi dan ketepatan model, dibuatlah visualisasi *confusion matrix* dan persebaran data berdasarkan kelas BMI. Grafik menunjukkan bahwa sebagian besar prediksi berada pada diagonal utama, yang menunjukkan bahwa model mampu mengidentifikasi kelas dengan benar.



Gambar 4.1 *Confusion Matrix*



Gambar 4.2 *Scatter Plot Tinggi vs Berat*



Gambar 4.3 *F1 Score*

4.3 Simulasi Prediksi dan Rekomendasi

Sistem juga menyediakan fitur input manual di mana pengguna dapat memasukkan data berupa jenis kelamin, tinggi, dan berat badan, lalu mendapatkan rekomendasi latihan berdasarkan hasil prediksi.

Contoh Simulasi:

- Input: Pria, 180 cm, 70 kg
- Prediksi: *Index* = 2 (Normal)
- Rekomendasi: *Maintenance* – Jaga pola makan & aktif bergerak

Pendekatan ini tidak hanya mempermudah pengguna dalam memahami kondisi tubuhnya, tetapi juga memfasilitasi pengambilan keputusan mandiri berbasis data. Sistem ini dapat dijadikan prototipe awal untuk pengembangan aplikasi kebugaran yang lebih personal dan interaktif.

4.4 Perbandingan dengan Penelitian Terkait

Penelitian ini memperoleh akurasi 92%, lebih tinggi dari hasil yang diperoleh Lestari dkk. yang hanya mencapai 68% pada klasifikasi status gizi anak disabilitas [1]. Hal ini kemungkinan disebabkan oleh dataset yang lebih bersih dan variabel yang lebih sederhana.

Aptana dkk. mencatat akurasi 96% dengan penerapan *SMOTE-ENN* pada klasifikasi BMI, menunjukkan bahwa penanganan ketidakseimbangan data dapat mendorong peningkatan performa lebih lanjut [10].

5. Kesimpulan dan Saran

5.1 Kesimpulan

Penelitian ini telah berhasil membangun sebuah sistem klasifikasi kategori BMI dan rekomendasi latihan fisik berbasis algoritma *K-Nearest Neighbors* (KNN). Dengan

menggunakan data BMI yang mencakup atribut jenis kelamin, tinggi badan, dan berat badan, sistem mampu menghasilkan akurasi sebesar 92% dalam mengklasifikasikan status gizi ke dalam enam kategori.

Setelah klasifikasi dilakukan, sistem memberikan rekomendasi latihan fisik berupa *bulking*, *maintenance*, atau *cutting* berdasarkan hasil klasifikasi. Hal ini menunjukkan bahwa metode *K-Nearest Neighbors* (KNN) efektif digunakan dalam membangun sistem berbasis data untuk mendukung keputusan kebugaran secara personal.

5.2 Saran

Untuk pengembangan lebih lanjut, disarankan:

- Menggunakan dataset yang lebih besar dan bervariasi untuk meningkatkan akurasi model.
- Menerapkan metode resampling seperti *SMOTE* untuk mengatasi ketidakseimbangan kelas.
- Mengembangkan sistem ke dalam bentuk aplikasi berbasis web atau mobile untuk meningkatkan aksesibilitas pengguna.
- Menambahkan fitur pelacakan progres pengguna secara berkala.

Penelitian ini juga dapat diperluas dengan membandingkan performa *K-Nearest Neighbors* (KNN) dengan algoritma lain seperti *Random Forest*, *SVM*, atau *Artificial Neural Network* untuk mendapatkan pendekatan terbaik dalam klasifikasi data kesehatan.

Referensi

- [1] E. Triwira Lestari and J. Adhiva, 'SENTIMAS: Seminar Nasional Penelitian dan Pengabdian Masyarakat Implementation Naive Bayes Classifier Algorithm and K-Nearest Neighbor For Obesity Nutritional Status of Children with Disabilities Implementasi Algoritma Naive Bayes Classifier dan K-Nearest Neighbor Untuk Klasifikasi Status Gizi Obesitas Anak Disabilitas', 2022, [Online]. Available: <https://journal.irpi.or.id/index.php/sentimas>
- [2] Z. Rizqika Khairunnisa *et al.*, 'Hubungan Asupan Zat Gizi Makro dan Aktivitas Fisik dengan Indeks Massa Tubuh (IMT) pada Mahasiswa Info Articles', *Journal of Biomedical Science and Health*, vol. 2, no. 2, pp. 9–13, 2022, doi: 10.31331/IJBSH.v2i1i1.2364.
- [3] M. R. Andreanto and A. Haryanto, 'JOSSAE (Journal of Sport Science and Education) Analisis Indeks Massa Tubuh dan Kondisi Fisik Atlet Floorball Kota Surabaya', *Journal of Sport Science and Education* |, vol. 6, no. 2, pp. 125–133, 2021, doi: 10.26740/j.
- [4] S. Hardwis, 'Analisis Resiko Obesitas Berdasarkan Aktivitas Fisik: Implementasi Metode Artificial Intelligence Machine Learning', 2024. [Online]. Available: <https://jurnal.unigal.ac.id/index.php/JKP>
- [5] A. Nikmatul Kasanah, U. Pujiyanto, T. Elektro, F. Teknik, and U. Negeri Malang, 'Terakreditasi SINTA Peringkat 2 Penerapan Teknik SMOTE untuk Mengatasi Imbalance Class dalam Klasifikasi Objektivitas Berita Online Menggunakan Algoritma KNN', *masa berlaku mulai*, vol. 1, no. 3, pp. 196–201, 2017.

- [6] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmaddeni, and L. Efrizoni, 'Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia', *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 1, pp. 273–281, Jan. 2024, doi: 10.57152/malcom.v4i1.1085.
- [7] N. Yogi Aptana, A. Nur Ikhsan, W. Maulana Baihaqi, and C. Raras Ajeng Widiawati, 'Perbandingan Random Forest dan K-Nearest Neighbors untuk Klasifikasi Body Mass Index Menggunakan SMOTE-ENN untuk Mengatasi Ketidakseimbangan Data pada Analisis Kesehatan', vol. 16, no. 01, 2025, doi: 10.35970/infotekmesin.v16i1.2553.
- [8] · Book, 'DATA MINING K-Nearest Neighbor', 2017. [Online]. Available: <https://www.researchgate.net/publication/321804055>
- [9] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, 'Text classification algorithms: A survey', 2019, *MDPI AG*. doi: 10.3390/info10040150.
- [10] M. Iqbal Nurkholis, D. Prabowo, B. Hari Prasetyo, S. Informasi, T. Informatika, and U. Muhammadiyah Metro, 'PERANCANGAN SISTEM INFORMASI PENERIMAAN PESERTA DIDIK BARU BERBASIS WEB PADA LKP KOMIHWAWA METRO SELATAN', 2024.