PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* UNTUK REKOMENDASI PROGRAM LATIHAN *WORKOUT* DI RUMAH BERDASARKAN KARAKTERISTIK PENGGUNA

Rico Herdiyanto[1*], Agus Junaidi[2]

[1,2] *Universitas Bina Sarana Informatika*
E-mail : herdiyantorico@gmail.com

ABSTRACT

The development of technology provides opportunities for the use of recommendation systems to assist users in determining workout programs according to their characteristics. This study aims to implement the K-Nearest Neighbor (KNN) algorithm to develop a home workout recommendation system based on user characteristics. This study employed a quantitative approach and machine learning using a dataset consisting of 973 records. The user characteristics used as input attributes included age, Body Mass Index (BMI), workout goals, workout experience, and available workout time. The dataset was divided into 80% training data and 20% testing data, followed by encoding and standardization using StandardScaler. K values ranging from 1 to 20 were tested to determine the optimal parameter. The results showed that K = 2 was the optimal value, achieving an accuracy of 97.44%. Validation using WEKA with the IBk algorithm resulted in an accuracy of 97.4359%, with 190 data correctly classified and 5 data misclassified. The system generated five workout recommendation categories: Full Body, HIIT, Cardio, Split Workout, and Push Pull Legs. These results indicate that the KNN algorithm can provide highly accurate workout recommendation classification and can be implemented in a web-based recommendation system.

Keywords : *K-Nearest Neighbor, recommendation system, workout, machine learning, Streamlit*

ABSTRAK

Perkembangan teknologi memberikan peluang dalam pemanfaatan sistem rekomendasi untuk membantu pengguna menentukan program latihan yang sesuai dengan karakteristiknya. Penelitian ini bertujuan menerapkan algoritma *K-Nearest Neighbor* (KNN) untuk membangun sistem rekomendasi *workout* di rumah berdasarkan karakteristik pengguna. Penelitian menggunakan pendekatan kuantitatif dan *machine learning* dengan dataset sebanyak 973 data. Karakteristik pengguna yang digunakan sebagai atribut masukan meliputi umur, *Body Mass Index* (BMI), tujuan latihan, pengalaman *workout*, dan waktu latihan yang tersedia. Data dibagi menjadi 80% data *training* dan 20% data *testing*, kemudian dilakukan *encoding* dan standardisasi menggunakan *StandardScaler*. Pengujian nilai K dilakukan pada rentang K = 1–20 untuk memperoleh parameter terbaik. Hasil penelitian menunjukkan bahwa nilai K optimal adalah K = 2 dengan akurasi sebesar 97,44%. Validasi menggunakan WEKA melalui algoritma IBk menghasilkan akurasi sebesar 97,4359%, dengan 190 data diklasifikasikan benar dan 5 data mengalami kesalahan klasifikasi. Sistem menghasilkan lima kategori rekomendasi, yaitu *Full Body*, *HIIT*, *Cardio*, *Split Workout*, dan *Push Pull Legs*. Hasil penelitian menunjukkan bahwa algoritma KNN mampu memberikan klasifikasi rekomendasi *workout* dengan tingkat akurasi yang tinggi dan dapat diterapkan pada sistem rekomendasi berbasis *web*.

Kata Kunci : *K-Nearest Neighbor, Sistem Rekomendasi, Workout, Machine Learning, Streamlit*

PENDAHULUAN

Olahraga merupakan salah satu aktivitas penting dalam menjaga kesehatan dan meningkatkan kapasitas fisik. Latihan fisik dapat meningkatkan kapasitas fungsional pada individu dengan hipertensi, diabetes melitus tipe 2, dan penyakit kardiovaskular, meskipun efektivitas latihan dapat berbeda berdasarkan kondisi individu dan jenis latihan yang dilakukan (Rijal et al., 2024). Pemilihan olahraga yang tidak sesuai dengan kondisi kesehatan juga berisiko menimbulkan cedera atau memperburuk kondisi medis yang telah ada (Kemalasari Siregar et al., 2025). Salah satu aktivitas yang banyak dilakukan untuk meningkatkan kebugaran adalah *fitness*, yang dapat ditujukan untuk meningkatkan massa otot, pembentukan otot, menurunkan maupun menambah berat badan (Rahmatullah & Adimahendra, 2022; Mubarak et al., 2021).

Perkembangan gaya hidup modern turut meningkatkan permasalahan kesehatan seperti obesitas, diabetes, dan penyakit jantung yang berkaitan dengan pola makan tidak sehat dan kurangnya aktivitas fisik. Rendahnya pemahaman mengenai pola makan sehat, pentingnya olahraga, serta faktor risiko kesehatan menjadi salah satu hambatan dalam upaya pencegahan (Irfan Sazali Nasution et al., 2025). Dalam kondisi tersebut, pengguna membutuhkan program latihan yang sesuai dengan karakteristik dan tujuan masing-masing. Namun, pemilihan program *workout* masih sering dilakukan berdasarkan tren media sosial atau mengikuti latihan orang lain tanpa mempertimbangkan usia, indeks massa tubuh (*Body Mass Index*/BMI), pengalaman latihan, tujuan olahraga, dan waktu latihan. Ketidaksesuaian program dapat menyebabkan hasil latihan kurang optimal serta meningkatkan risiko cedera dan kelelahan. Hanani et al. (2024) menunjukkan bahwa pengalaman berlatih memiliki pengaruh terhadap risiko cedera, sehingga karakteristik pengguna penting dipertimbangkan dalam pemilihan program latihan.

Salah satu pendekatan yang dapat digunakan untuk membantu permasalahan tersebut adalah sistem rekomendasi. Sistem rekomendasi mampu memberikan saran berdasarkan data dan karakteristik pengguna sehingga dapat menghasilkan rekomendasi yang lebih personal (Roy & Dutta, 2022). Salah satu metode *machine learning* yang dapat diterapkan adalah algoritma *K-Nearest Neighbor* (KNN), yaitu metode yang menentukan hasil berdasarkan kedekatan suatu data dengan sejumlah data terdekat. KNN memiliki karakteristik sederhana dan mudah diimplementasikan serta telah digunakan dalam berbagai permasalahan klasifikasi dan rekomendasi (Ning et al., 2015). Dalam bidang olahraga, KNN telah digunakan untuk memberikan rekomendasi aktivitas fisik berdasarkan karakteristik pengguna dan dilaporkan mampu mencapai akurasi hingga 90% (Maulyanda et al., 2025). Wibowo dan Erawati (2025) juga menunjukkan bahwa KNN memiliki performa yang baik dalam memprediksi aktivitas olahraga berdasarkan karakteristik pengguna.

Berdasarkan permasalahan tersebut, penelitian ini menerapkan algoritma *K-Nearest Neighbor* untuk membangun sistem rekomendasi program latihan *workout* di rumah berdasarkan karakteristik pengguna. Sistem dikembangkan dalam bentuk *website* interaktif menggunakan Streamlit sehingga pengguna dapat memperoleh rekomendasi latihan secara lebih mudah dan sesuai dengan kondisi serta tujuan latihan. Penelitian ini bertujuan untuk menerapkan KNN, membangun sistem rekomendasi *workout*, serta mengukur tingkat akurasi algoritma dalam menghasilkan rekomendasi program latihan.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *machine learning* untuk membangun sistem rekomendasi *workout* di rumah berbasis *web*. Algoritma yang digunakan adalah *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan jenis *workout* berdasarkan karakteristik pengguna, yaitu umur, *Body Mass Index* (BMI), tujuan latihan, tingkat pengalaman, dan durasi latihan. Sistem dikembangkan menggunakan bahasa pemrograman Python, *framework* Streamlit, dan pustaka *Scikit-Learn*.

Data yang digunakan berupa dataset sebanyak 973 data yang dikembangkan berdasarkan *Gym Members Exercise Dataset* dari Kaggle. Dataset asli digunakan sebagai referensi dalam menentukan atribut yang relevan, kemudian disesuaikan dengan kebutuhan penelitian. Atribut yang digunakan meliputi umur, BMI, tujuan latihan, pengalaman, waktu latihan, dan rekomendasi *workout*. Label rekomendasi terdiri atas *Full Body*, *HIIT*, *Cardio*, *Split Workout*, dan *Push Pull Legs*. Dataset disimpan dalam format CSV dan dibagi menggunakan metode *hold-out validation* dengan perbandingan 80% data *training* (778 data) dan 20% data *testing* (195 data).

Tahapan penelitian meliputi pengumpulan data, *preprocessing*, implementasi KNN, pengembangan sistem, dan pengujian. Pada tahap *preprocessing* dilakukan pemeriksaan data, *encoding* terhadap atribut kategorikal menggunakan *LabelEncoder*, serta standardisasi data menggunakan *StandardScaler*. Selanjutnya, data dipisahkan menjadi variabel fitur yang terdiri atas umur, BMI, tujuan, pengalaman, dan waktu, sedangkan variabel target adalah rekomendasi *workout*.

Algoritma KNN diterapkan dengan menghitung jarak antara data pengguna dan data *training* menggunakan *Euclidean Distance*. Data kemudian diurutkan berdasarkan jarak terdekat dan sejumlah K tetangga digunakan untuk menentukan kelas rekomendasi berdasarkan mayoritas. Pengujian nilai K dilakukan pada rentang K = 1 hingga K = 20. Nilai K terbaik ditentukan berdasarkan akurasi tertinggi pada data *testing*. Berdasarkan hasil pengujian, nilai K = 2 menghasilkan akurasi tertinggi sebesar 97,44% dan digunakan sebagai parameter model akhir.

Sistem rekomendasi dikembangkan menggunakan Streamlit sebagai antarmuka berbasis *web*. Pengguna memasukkan umur, BMI, tujuan latihan, tingkat pengalaman, dan waktu latihan. Data masukan kemudian melalui proses *encoding* dan standardisasi sebelum diproses oleh model KNN untuk menghasilkan rekomendasi *workout*.

Evaluasi model dilakukan menggunakan metrik *accuracy* untuk mengetahui kemampuan KNN dalam mengklasifikasikan data *testing*. Akurasi dihitung berdasarkan perbandingan jumlah prediksi yang benar dengan keseluruhan data pengujian. Selain evaluasi model, sistem diuji berdasarkan kemampuannya menerima input pengguna, menjalankan proses prediksi, dan menampilkan rekomendasi *workout* melalui antarmuka Streamlit (*Gym Members Exercise Dataset*, Kaggle).

HASIL PENELITIAN DAN PEMBAHASAN

Hasil

Penelitian ini menghasilkan sistem rekomendasi *workout* berbasis *web* menggunakan algoritma *K-Nearest Neighbor* (KNN). Sistem dikembangkan menggunakan Python dan *framework* Streamlit dengan lima kelas rekomendasi, yaitu *Full Body*, *HIIT*, *Cardio*, *Split*

Workout, dan *Push Pull Legs*. Rekomendasi diberikan berdasarkan karakteristik pengguna yang meliputi umur, *Body Mass Index* (BMI), tujuan latihan, tingkat pengalaman, dan waktu latihan.

Dataset Penelitian

Dataset yang digunakan terdiri dari 973 data dengan enam atribut, yaitu umur, BMI, tujuan latihan, pengalaman, waktu latihan, dan rekomendasi. Dataset dikembangkan berdasarkan *Gym Members Exercise Dataset* dari Kaggle dan disesuaikan dengan kebutuhan sistem rekomendasi.

Sebagian data penelitian yang digunakan ditampilkan pada Tabel 1.

Tabel 1. Contoh Data Penelitian

Umur	BMI	Tujuan	Pengalaman	Waktu	Rekomendasi
22	22	*muscle_gain*	*beginner*	45	*Full Body*
25	30	*fat_loss*	*beginner*	30	*HIIT*
28	24	*endurance*	*intermediate*	60	*Cardio*
30	26	*muscle_gain*	*advanced*	75	*Push Pull Legs*
21	21	*muscle_gain*	*beginner*	60	*Split Workout*

Tabel 1 menunjukkan contoh data yang digunakan dengan kombinasi karakteristik pengguna dan label rekomendasi *workout*. Data tersebut menjadi dasar bagi model KNN dalam mempelajari pola hubungan antara karakteristik pengguna dan jenis latihan.

Sistem memiliki lima kelas keluaran yang digunakan dalam proses klasifikasi. Kategori dan deskripsinya disajikan pada Tabel 2.

Tabel 2. Kelas Output Sistem

No.	*Workout*	Deskripsi
1	*Full Body*	Latihan seluruh kelompok otot dalam satu sesi
2	*HIIT*	Latihan intensitas tinggi dengan interval
3	*Cardio*	Latihan untuk meningkatkan daya tahan dan pembakaran kalori
4	*Split Workout*	Latihan dengan pembagian kelompok otot ke beberapa sesi
5	*Push Pull Legs*	Latihan berdasarkan pola gerakan *push*, *pull*, dan *legs*

Tabel 2 memperlihatkan lima kelas yang menjadi target prediksi model. Setiap kelas mewakili jenis program latihan yang dapat direkomendasikan kepada pengguna berdasarkan karakteristik input.

Implementasi Sistem

Sistem yang dikembangkan menyediakan antarmuka berbasis *web* untuk memasukkan karakteristik pengguna dan memperoleh hasil rekomendasi. Halaman utama menampilkan informasi sistem, nilai K terbaik, dan akurasi model.



Gambar 1. Halaman Utama Sistem

Gambar 1 menunjukkan halaman utama sistem yang menjadi tampilan awal aplikasi. Informasi mengenai model dan hasil pengujian ditampilkan untuk memberikan gambaran mengenai sistem rekomendasi yang digunakan.

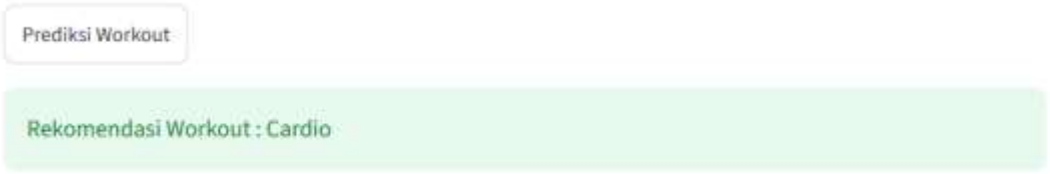
Pengguna selanjutnya dapat memasukkan data sesuai karakteristik yang diperlukan oleh model.



Gambar 2. Halaman Input Sistem

Gambar 2 memperlihatkan halaman input yang digunakan untuk memasukkan umur, BMI, tujuan latihan, pengalaman, dan waktu latihan. Data tersebut menjadi masukan bagi proses prediksi KNN.

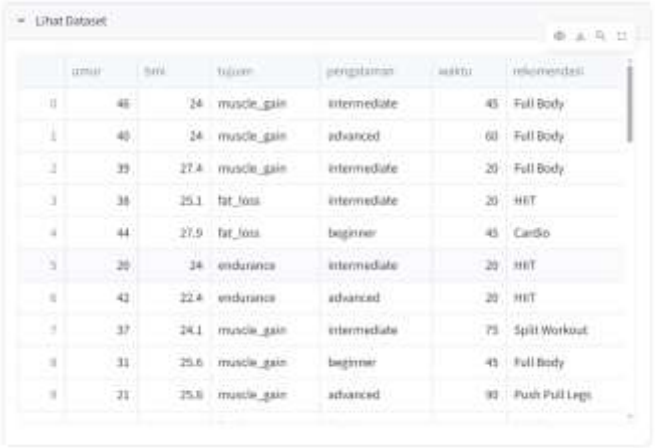
Setelah data dimasukkan, sistem melakukan proses prediksi dan menampilkan jenis *workout* yang direkomendasikan.



Gambar 3. Halaman Hasil Prediksi

Gambar 3 menunjukkan hasil prediksi sistem berdasarkan data pengguna. Hasil yang ditampilkan berupa rekomendasi jenis *workout* yang diperoleh dari model KNN.

Selain halaman input dan hasil prediksi, sistem menyediakan tampilan sebagian dataset yang digunakan.



Gambar 4. Halaman Dataset *Workout*

Gambar 4 memperlihatkan fitur tampilan dataset pada sistem. Fitur tersebut digunakan untuk melihat data yang menjadi dasar pelatihan model.

Pengujian Model KNN

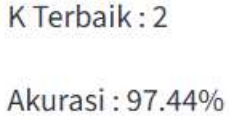
Pengujian dilakukan untuk menentukan nilai K yang memberikan performa klasifikasi terbaik. Pengujian dilakukan terhadap nilai K = 1 sampai K = 20 menggunakan data *testing*.

Tabel 3. Hasil Pengujian Nilai K

Nilai K	Akurasi (%)
1	95,90
2	97,44
3	96,92
4	95,90
5	96,41
6	92,82
7	94,36
8	90,77
9	91,79
10	90,77
11	91,79
12	88,72
13	88,21
14	88,21
15	90,26
16	90,26
17	89,74
18	89,23
19	90,26
20	89,23

Berdasarkan Tabel 3, akurasi tertinggi diperoleh pada K = 2 dengan nilai 97,44%. Nilai tersebut kemudian digunakan sebagai parameter model akhir. Hasil ini menunjukkan bahwa penggunaan dua tetangga terdekat memberikan performa klasifikasi paling baik dibandingkan nilai K lainnya yang diuji.

Hasil akurasi model kemudian ditampilkan pada sistem seperti pada Gambar 5.



Gambar 5. Hasil Akurasi Model KNN

Gambar 5 menunjukkan bahwa model KNN memperoleh akurasi sebesar 97,44%. Nilai tersebut menunjukkan tingkat ketepatan model dalam mengklasifikasikan data *testing* berdasarkan lima kelas rekomendasi *workout*.

Pengujian Prediksi Sistem

Pengujian prediksi dilakukan menggunakan beberapa skenario karakteristik pengguna untuk melihat kemampuan sistem dalam menghasilkan rekomendasi. Pada pengujian pertama, pengguna memiliki tujuan *muscle gain*, pengalaman *beginner*, dan waktu latihan relatif terbatas.



Gambar 6. Hasil Pengujian Prediksi 1

Gambar 6 menunjukkan sistem menghasilkan rekomendasi *Full Body*. Rekomendasi tersebut merupakan hasil klasifikasi berdasarkan kombinasi karakteristik pengguna yang dimasukkan ke sistem.

Pada pengujian kedua, pengguna memiliki BMI yang cukup tinggi dan tujuan latihan *fat loss*.



Gambar 7. Hasil Pengujian Prediksi 2

Gambar 7 menunjukkan sistem memberikan rekomendasi *HIIT*. Hasil tersebut diperoleh berdasarkan karakteristik input yang diberikan pada sistem.

Pengujian ketiga menggunakan pengguna dengan tujuan meningkatkan daya tahan atau *endurance*.



Gambar 8. Hasil Pengujian Prediksi 3

Gambar 8 menunjukkan rekomendasi *Cardio* yang dihasilkan sistem untuk karakteristik pengguna dengan tujuan meningkatkan daya tahan tubuh.

Pada pengujian keempat, pengguna memiliki tingkat pengalaman *advanced* dan waktu latihan yang cukup panjang.



Gambar 9. Hasil Pengujian Prediksi 4

Gambar 9 menunjukkan sistem memberikan rekomendasi *Push Pull Legs*. Hasil ini merupakan keluaran model berdasarkan karakteristik pengguna yang dimasukkan.

Pada pengujian kelima, pengguna memiliki tujuan *muscle gain* dengan waktu latihan yang cukup.



Gambar 10. Hasil Pengujian Prediksi 5

Gambar 10 menunjukkan sistem memberikan rekomendasi *Split Workout*. Hasil prediksi menunjukkan bahwa model dapat memberikan kelas rekomendasi yang berbeda sesuai dengan kombinasi karakteristik input pengguna.

Sebagai gambaran tambahan mengenai hubungan tujuan latihan dengan jenis latihan, hasil rekomendasi berdasarkan target ditampilkan pada Tabel 4.

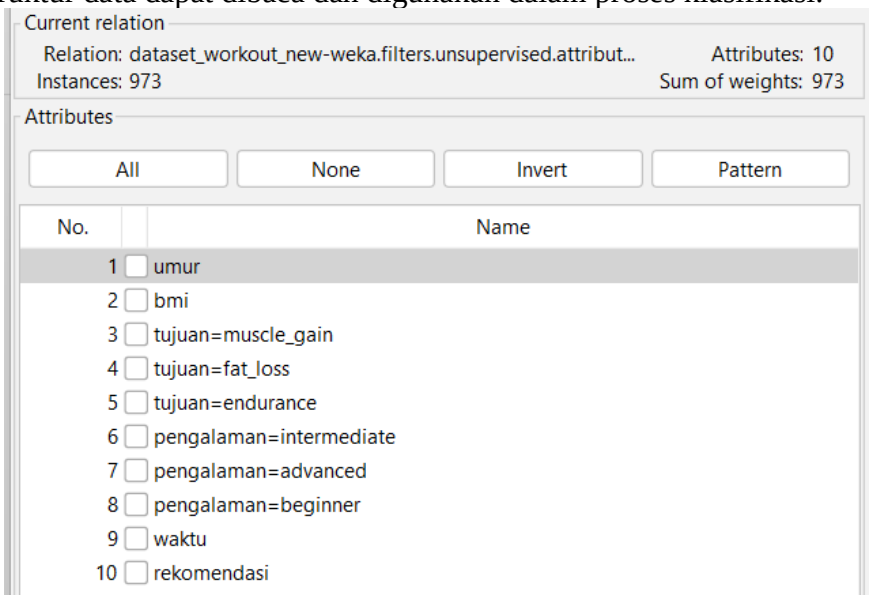
Tabel 4. Rekomendasi *Workout* Berdasarkan Target

Target/Tujuan	Jenis Latihan	Frekuensi	Durasi	Contoh Gerakan/Intensitas
Menurunkan berat badan (*Fat Loss*)	*Cardio & HIIT*	3–5×/minggu	30–45 menit	*Jumping jacks*, *burpees*, *mountain climbers*, *brisk walking*, *cycling interval*
Membentuk otot (*Hypertrophy*)	Latihan beban (*Resistance Training*)	3–4×/minggu	45–60 menit	*Squat*, *push-up*, *deadlift*, *bench press*, *dumbbell row* (70–80% 1RM)
Meningkatkan kebugaran dan stamina	Kombinasi *Cardio & Calisthenics*	3×/minggu	30–60 menit	Lari jarak menengah, *rowing*, *bodyweight workout circuit*
Fleksibilitas dan pemulihan (*Recovery*)	Latihan intensitas rendah (*Low Intensity*)	1–2×/minggu	20–30 menit	*Yoga*, *Pilates*, *dynamic stretching*, *foam rolling*

Tabel 4 memberikan gambaran umum keterkaitan antara target latihan, jenis aktivitas, frekuensi, durasi, serta contoh gerakan. Tabel ini mendukung konteks rekomendasi latihan, sedangkan rekomendasi utama dalam sistem ditentukan berdasarkan hasil klasifikasi KNN.

Validasi Menggunakan WEKA

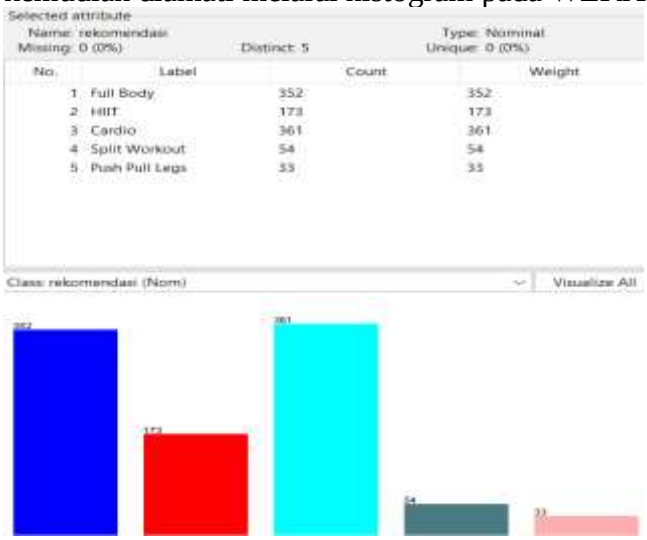
Validasi tambahan dilakukan menggunakan WEKA melalui algoritma IBk (*Instance-Based K Nearest Neighbor*). Dataset dimuat melalui menu *Explorer* pada bagian *Preprocess* untuk memastikan struktur data dapat dibaca dan digunakan dalam proses klasifikasi.



Gambar 11. Tampilan Dataset pada WEKA

Gambar 11 menunjukkan dataset berhasil dimuat ke dalam WEKA dengan 973 *instances* dan enam atribut, yaitu umur, BMI, tujuan latihan, pengalaman, waktu latihan, dan rekomendasi *workout*. Atribut rekomendasi digunakan sebagai *class attribute*.

Distribusi kelas kemudian diamati melalui histogram pada WEKA.



Gambar 12. Histogram Atribut Rekomendasi

Gambar 12 menunjukkan terdapat lima kelas rekomendasi, yaitu *Full Body*, *HIIT*, *Cardio*, *Split Workout*, dan *Push Pull Legs*. Jumlah data pada setiap kelas berbeda, sehingga distribusi kelas perlu diperhatikan dalam interpretasi hasil klasifikasi.

Selanjutnya dilakukan pengujian menggunakan IBk pada menu *Classify*.

```
=== Summary ===

Correctly Classified Instances         190               97.4359 %
Incorrectly Classified Instances         5                2.5641 %
Kappa statistic                          0.9634
Mean absolute error                      0.0187
Root mean squared error                  0.1007
Relative absolute error                  6.6966 %
Root relative squared error             26.8266 %
Total Number of Instances              195

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               1.000    0.008    0.985      1.000    0.992      0.988    1.000     1.000     Full Bod
               0.976    0.019    0.930      0.976    0.952      0.940    0.995     0.962     HIIT
               0.959    0.008    0.986      0.959    0.973      0.956    0.991     0.988     Cardio
               0.889    0.000    1.000      0.889    0.941      0.940    0.999     0.989     Split Wo
               1.000    0.000    1.000      1.000    1.000      1.000    1.000     1.000     Push Pul
Weighted Avg.  0.974    0.010    0.975      0.974    0.974      0.964    0.996     0.987
```

Gambar 13. Hasil Pengujian IBk (KNN)

Hasil pengujian menunjukkan bahwa dari 195 data *testing*, sebanyak 190 data berhasil diklasifikasikan dengan benar dan lima data mengalami kesalahan klasifikasi. Dengan demikian, akurasi yang diperoleh adalah 97,4359%, sedangkan *error rate* sebesar 2,5641%. Hasil validasi WEKA tersebut sangat dekat dengan akurasi pengujian menggunakan Python dan Streamlit sebesar 97,44%.

Analisis *Confusion Matrix*

Untuk melihat hasil klasifikasi pada masing-masing kelas secara lebih rinci, dilakukan analisis *Confusion Matrix* dari hasil pengujian IBk.

```
=== Confusion Matrix ===

  a  b  c  d  e   <-- classified as
 64  0  0  0  0 |  a = Full Body
  1 40  0  0  0 |  b = HIIT
  0  3 71  0  0 |  c = Cardio
  0  0  1  8  0 |  d = Split Workout
  0  0  0  0  7 |  e = Push Pull Legs
```

Gambar 14. *Confusion Matrix* Pengujian Ibk

Berdasarkan *Confusion Matrix*, sebanyak 46 data kelas *Full Body* berhasil diklasifikasikan dengan benar tanpa kesalahan. Pada kelas *HIIT*, 40 data diklasifikasikan dengan benar dan satu data diklasifikasikan sebagai *Full Body*. Kelas *Cardio* memiliki 71 data yang diklasifikasikan dengan benar, sedangkan tiga data diklasifikasikan sebagai *HIIT*. Pada kelas *Split Workout*, delapan data diklasifikasikan dengan benar dan satu data diklasifikasikan sebagai *Cardio*. Sementara itu, tujuh data *Push Pull Legs* berhasil diklasifikasikan dengan benar tanpa kesalahan.

Secara keseluruhan, *Confusion Matrix* menunjukkan bahwa sebagian besar data berhasil diklasifikasikan sesuai kelas sebenarnya. Kesalahan klasifikasi terutama terjadi antara kelas

Cardio dan *HIIT*, yang menunjukkan adanya kemiripan karakteristik pada data yang digunakan. Hasil validasi tersebut memperkuat hasil pengujian model pada sistem, dengan akurasi sekitar 97,44%, sehingga KNN dapat digunakan untuk menghasilkan rekomendasi *workout* berdasarkan karakteristik pengguna.

Pembahasan

Analisis Hasil Implementasi

Hasil penelitian menunjukkan bahwa algoritma *K-Nearest Neighbor* (KNN) dapat diterapkan pada sistem rekomendasi *workout* berbasis *web* menggunakan Python dan Streamlit. Model menggunakan umur, *Body Mass Index* (BMI), tujuan latihan, pengalaman, dan waktu latihan sebagai karakteristik pengguna. KNN menentukan rekomendasi berdasarkan kedekatan karakteristik data pengguna dengan data *training*, sehingga data dengan kemiripan lebih tinggi menjadi dasar penentuan kelas *workout*. Penggunaan KNN untuk sistem rekomendasi sejalan dengan Ning et al. (2015) yang menunjukkan bahwa KNN dapat digunakan dalam proses klasifikasi berdasarkan kemiripan karakteristik data.

Hasil implementasi menunjukkan sistem mampu menghasilkan lima jenis rekomendasi, yaitu *Full Body*, *HIIT*, *Cardio*, *Split Workout*, dan *Push Pull Legs*. Hal tersebut menunjukkan bahwa atribut yang digunakan dapat dimanfaatkan untuk membedakan karakteristik pengguna dalam menentukan rekomendasi *workout*. Temuan ini sejalan dengan Maulyanda et al. (2025) dan Wibowo dan Erawati (2025) yang menunjukkan penerapan KNN dalam memberikan rekomendasi atau prediksi aktivitas olahraga berdasarkan karakteristik pengguna.

Analisis Hasil Pengujian Model

Pengujian nilai K = 1 sampai K = 20 menghasilkan akurasi tertinggi sebesar 97,44% pada K = 2. Hasil tersebut menunjukkan bahwa pemilihan nilai K mempengaruhi performa KNN. Meskipun K = 1 juga menghasilkan akurasi tinggi sebesar 95,90%, K = 2 dipilih karena memberikan akurasi terbaik pada data *testing*. Hasil akurasi yang tinggi menunjukkan bahwa model mampu mengklasifikasikan sebagian besar data pengujian sesuai dengan kelas rekomendasi.

Hasil tersebut memperlihatkan bahwa KNN dapat memberikan performa klasifikasi yang baik ketika karakteristik data pengguna memiliki pola yang cukup jelas. Temuan ini mendukung penelitian Maulyanda et al. (2025) dan Wibowo dan Erawati (2025) yang menunjukkan bahwa metode KNN dapat digunakan untuk permasalahan yang berkaitan dengan rekomendasi aktivitas berdasarkan karakteristik tertentu. Namun, tingginya akurasi perlu dipahami sebagai performa klasifikasi terhadap dataset yang digunakan dan belum secara langsung menunjukkan bahwa rekomendasi tersebut merupakan program latihan terbaik secara medis bagi setiap individu.

Validasi Menggunakan WEKA

Validasi menggunakan WEKA melalui algoritma IBk menghasilkan 190 data yang diklasifikasikan dengan benar dan lima data yang salah dari 195 data pengujian, dengan akurasi sebesar 97,4359%. Hasil tersebut sangat dekat dengan akurasi implementasi Python sebesar 97,44%, sehingga memberikan dukungan tambahan terhadap hasil pengujian model.

Perbedaan kecil pada angka akurasi disebabkan oleh pembulatan, sehingga kedua hasil dapat dinyatakan memiliki performa yang hampir sama. Hasil validasi ini memperkuat bahwa model KNN mampu melakukan klasifikasi terhadap data rekomendasi *workout* dengan tingkat ketepatan yang tinggi. Penggunaan algoritma IBk pada WEKA juga sesuai dengan konsep KNN yang melakukan klasifikasi berdasarkan kedekatan atau kemiripan data (Ning et al., 2015).

Analisis *Confusion Matrix*

Hasil *confusion matrix* menunjukkan bahwa sebagian besar data pada setiap kelas berhasil diklasifikasikan dengan benar. Kelas *Full Body* memiliki 46 prediksi benar, *HIIT* sebanyak 40, *Cardio* sebanyak 71, *Split Workout* sebanyak 8, dan *Push Pull Legs* sebanyak 7. Kesalahan klasifikasi ditemukan pada beberapa kelas, terutama *Cardio* yang sebanyak tiga datanya diprediksi sebagai *HIIT*.

Kesalahan tersebut menunjukkan adanya kemiripan karakteristik antara beberapa kelas *workout*. *Cardio* dan *HIIT*, misalnya, dapat memiliki pola tujuan dan durasi latihan yang berdekatan pada sebagian data. Kondisi tersebut dapat menyebabkan jarak antar-data menjadi relatif dekat sehingga KNN menghasilkan klasifikasi yang berbeda dari label sebenarnya. Kondisi ini sesuai dengan prinsip KNN yang menentukan kelas berdasarkan kedekatan data dengan sejumlah *neighbor* terdekat (Ning et al., 2015).

Kelebihan dan Keterbatasan Sistem

Sistem memiliki kelebihan berupa kemampuan memberikan rekomendasi secara otomatis, proses prediksi yang sederhana, serta antarmuka *web* yang mudah digunakan melalui Streamlit. Penggunaan KNN juga memungkinkan proses klasifikasi berdasarkan kemiripan karakteristik pengguna tanpa perhitungan manual. Hasil tersebut mendukung penelitian Maulyanda et al. (2025) dan Wibowo dan Erawati (2025) mengenai penggunaan KNN untuk mendukung rekomendasi atau prediksi aktivitas olahraga.

Di sisi lain, penelitian memiliki beberapa keterbatasan. Dataset yang digunakan hanya terdiri dari 973 data dan sistem menggunakan lima atribut utama sebagai dasar rekomendasi. Jenis *workout* juga masih terbatas pada lima kategori. Selain itu, sistem belum mempertimbangkan faktor kesehatan seperti riwayat cedera dan kondisi medis tertentu. Dengan demikian, hasil rekomendasi lebih tepat digunakan sebagai alat bantu pemilihan latihan, bukan sebagai pengganti penilaian profesional.

Ketercapaian Tujuan Penelitian

Berdasarkan hasil implementasi dan pengujian, tujuan penelitian telah tercapai dengan berhasil dibangunnya sistem rekomendasi *workout* berbasis KNN menggunakan Streamlit. Sistem dapat memberikan rekomendasi berdasarkan umur, BMI, tujuan latihan, pengalaman, dan waktu latihan. Model memperoleh akurasi 97,44% pada implementasi Python dan 97,4359% pada validasi WEKA. Hasil tersebut menunjukkan bahwa KNN memiliki performa klasifikasi yang baik dan dapat diterapkan sebagai metode dalam sistem rekomendasi *workout* berbasis *web*, sebagaimana juga ditunjukkan dalam penelitian Maulyanda et al. (2025) dan Wibowo dan Erawati (2025).

KESIMPULAN

Penelitian ini berhasil menerapkan algoritma *K-Nearest Neighbor* (KNN) pada sistem rekomendasi *workout* berbasis *web* menggunakan Streamlit. Sistem memberikan rekomendasi berdasarkan umur, BMI, tujuan latihan, pengalaman, dan waktu latihan dengan lima kategori, yaitu *Full Body*, *HIIT*, *Cardio*, *Split Workout*, dan *Push Pull Legs*. Nilai K terbaik adalah K = 2 dengan akurasi 97,44%, sedangkan validasi menggunakan WEKA menghasilkan akurasi 97,4359%. Hasil tersebut menunjukkan bahwa KNN mampu mengklasifikasikan rekomendasi *workout* dengan baik dan dapat diterapkan sebagai metode rekomendasi latihan berbasis karakteristik pengguna.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam pelaksanaan penelitian ini, sehingga penelitian dapat diselesaikan dengan baik.

DAFTAR PUSTAKA

Hanani, A. D., Samat, S., Rosita, M., & Assapari, M. H. (2024). Profil cedera dan faktor risiko pada pengunjung Gym Numero Uno berdasarkan usia dan pengalaman berlatih. *Babul Ilmi Jurnal Ilmiah Multi Science Kesehatan, 16*(2). https://doi.org/10.36729/bi.v16i2.1294

Irfan Sazali Nasution, Arini Dwi Rahmadani, Audina, W., Sari, D. P., & Sari, N. D. (2025). Systematic review: Pengaruh gaya hidup dan pengetahuan masyarakat terhadap risiko penyakit jantung koroner. *Sehat Rakyat: Jurnal Kesehatan Masyarakat, 4*(2). https://doi.org/10.54259/sehatrakyat.v4i2.4337

Kemalasari Siregar, G., Pujianto, & Ridhawati, E. (2025). Sistem pendukung keputusan untuk rekomendasi olahraga berdasarkan kondisi kesehatan menggunakan metode *Analytical Hierarchy Process* (AHP). *Jurnal Ilmiah Sistem Informasi (JISI), 4*(1). https://doi.org/10.24127/jisi.v4i1.8561

Maulyanda, M., Deviani, R., Sabrina, F., & Afdhaluzzikri, A. (2025). WorkoutLife: Penerapan algoritma *K-Nearest Neighbor* (KNN) untuk rekomendasi workout berdasarkan data gaya hidup. *J-SIGN (Journal of Informatics, Information System, and Artificial Intelligence), 3*(2). https://doi.org/10.24815/j-sign.v3i2.50115

Mubarak, N., Kusumawati, A., & Raharjo, K. (2021). Menggali peran digital fitness influencer terhadap *parasocial interaction* dan *fan citizenship behaviour*. *Jurnal Studi Komunikasi (Indonesian Journal of Communications Studies), 5*(3). https://doi.org/10.25139/jsk.v5i3.3633

Ning, X., Desrosiers, C., & Karypis, G. (2015). A comprehensive survey of neighborhood-based recommendation methods. In *Recommender Systems Handbook* (2nd ed.). https://doi.org/10.1007/978-1-4899-7637-6_2

Rahmatullah, R., & Adimahendra, W. (2022). Sistem pendukung keputusan latihan gerakan fitness bagi pemula menggunakan metode *Weighted Product*. *Journal of Information System and Computer, 2*(2). https://doi.org/10.34001/jister.v2i2.420

Rijal, A., Adhikari, T. B., Dhakal, S., Maagaard, M., Piri, R., Nielsen, E. E., Neupane, D., Jakobsen, J. C., & Olsen, M. H. (2024). Effect of exercise on functional capacity and body weight for people with hypertension, type 2 diabetes, or cardiovascular disease: A systematic review with meta-analysis and trial sequential analysis. *BMC Sports Science, Medicine and Rehabilitation, 16*(1), Article 38. https://doi.org/10.1186/s13102-024-00829-1

Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data, 9*(1). https://doi.org/10.1186/s40537-022-00592-5

Valakhorasani, V. (n.d.). *Gym members exercise dataset* [Data set]. Kaggle. Kaggle

Wibowo, A., & Erawati, W. (2025). Pembuatan model rekomendasi gerakan olahraga berbasis *content-based filtering* dengan integrasi *clustering* dan klasifikasi. *Jurnal Infortech, 7*(2). https://doi.org/10.31294/infortech.v7i2.11384