Journal of System & Technology
ISSN: 3110-0171
Vol. 2, No. 1, June 2026

ANALISIS PERBANDINGAN ALGORITMA K-NEAREST NEIGHBORS DAN SUPPORT VECTOR MACHINE UNTUK KLASIFIKASI TINGKAT PENGALAMAN ANGGOTA GYM

Hilmi Aminuddien[1], Syachrul Ramadhan[1], Mohammad Ferry Irwansyah[1], Dian Ramadhani[1]*
[1]Teknik Informatika, Universitas Riau, Indonesia
hilmi.aminuddien2162@student.unri.ac.id, syachrul.ramadhan2354@student.unri.ac.id,
mohammad.ferry6145@student.unri.ac.id, dianramadhani@lecturer.unri.ac.id
*Corresponding author: dianramadhani@lecturer.unri.ac.id

Abstract: Penentuan tingkat pengalaman anggota gym secara akurat merupakan tantangan penting dalam perancangan program latihan yang personal dan efektif. Penelitian terdahulu yang membandingkan K-Nearest Neighbors (K-NN) dan Support Vector Machine (SVM) umumnya berfokus pada domain medis seperti prediksi diabetes, sehingga efektivitas kedua algoritma pada data kebugaran fisik belum banyak dikaji. Penelitian ini mengisi kesenjangan tersebut dengan menerapkan dan membandingkan K-NN dan SVM pada dataset Gym Members Exercise Tracking (973 data, 15 atribut) untuk mengklasifikasikan tingkat pengalaman anggota menjadi tiga kelas: Beginner, Intermediate, dan Expert. Pemrosesan data mencakup label encoding, normalisasi StandardScaler, serta pembagian data latih dan uji dengan rasio 80:20. K-NN diimplementasikan dengan K=5 menggunakan metrik jarak Euclidean, sedangkan SVM menggunakan kernel RBF dengan parameter C=1,0 dan gamma='scale'. Hasil pengujian menunjukkan SVM mengungguli K-NN dengan akurasi 85,64% (macro F1-score: 0,86) dibandingkan K-NN dengan akurasi 81,54% (macro F1-score: 0,81). Kontribusi utama penelitian ini adalah bukti empiris bahwa SVM dengan kernel RBF lebih efektif untuk klasifikasi multikelas pada data kebugaran fisik dibandingkan K-NN, serta identifikasi fitur-fitur fisiologis yang paling berpengaruh terhadap penentuan tingkat pengalaman anggota gym.

Keywords—Akurasi, Anggota Gym, K-NN, Klasifikasi, SVM

Article Information:
Received: 11-Jun-2026; Revised: 21-Jun-2026; Accepted: 26-Jun-2026; Published: 27-Jun-2026

This work is licensed under a **CC BY-SA**. Copyright ©2026 by Author. Published by Universitas Riau.

I. PENDAHULUAN

Kesehatan dan kebugaran fisik telah menjadi perhatian global yang terus berkembang. Pusat kebugaran (gym) menjadi salah satu sarana yang banyak dimanfaatkan masyarakat untuk meningkatkan kondisi fisik mereka. Setiap anggota gym memiliki tingkat pengalaman yang berbeda-beda, mulai dari pemula (Beginner), menengah (Intermediate), hingga mahir (Expert). Penentuan tingkat pengalaman ini sangat penting untuk merancang program latihan yang sesuai, aman, dan efektif bagi masing-masing individu[1].

Dalam beberapa tahun terakhir, pemanfaatan *machine learning* dalam bidang kesehatan dan kebugaran mengalami perkembangan yang signifikan [2]. Algoritma klasifikasi seperti *K-Nearest Neighbors* (K-NN) dan *Support Vector Machine* (SVM) telah banyak diaplikasikan dalam berbagai domain untuk mengenali pola aktivitas fisik dan memprediksi kondisi kebugaran [3]. K-NN bekerja berdasarkan prinsip kemiripan jarak antar data, sedangkan SVM membangun *hyperplane* optimal yang memisahkan kelas-kelas data secara maksimal [4].

Berbagai penelitian telah mengkaji efektivitas K-NN dan SVM dalam konteks klasifikasi. Mulyani et al. menunjukkan bahwa K-NN mampu mendeteksi pola data dengan baik namun performanya menurun pada dataset dengan fitur berskala beragam [3]. Penelitian lain pada domain serupa menunjukkan SVM cenderung lebih unggul pada data multikelas non-linear berkat kemampuan kernel mapping-nya [4]. Namun, hampir seluruh studi perbandingan yang ada berfokus pada dataset medis, khususnya prediksi penyakit diabetes [7][3], sehingga belum ada kajian yang secara spesifik mengevaluasi kedua algoritma pada data aktivitas fisik kebugaran (gym). Selain itu, penelitian terdahulu jarang menyertakan analisis pengaruh fitur (feature importance) terhadap keputusan klasifikasi, sehingga interpretabilitas model masih terbatas. Kesenjangan inilah yang menjadi motivasi utama penelitian ini untuk mengkaji perbandingan K-NN dan SVM secara empiris pada domain kebugaran fisik.

Berdasarkan kajian literatur tersebut, penelitian ini bertujuan untuk: (1) mengimplementasikan K-NN dan SVM pada dataset *Gym Members Exercise Tracking* untuk mengklasifikasikan tingkat pengalaman anggota gym menjadi kelas *Beginner*, *Intermediate*, dan *Expert*; (2) membandingkan kinerja kedua algoritma berdasarkan metrik akurasi, presisi, *recall*, dan F1-score; (3) mengidentifikasi fitur fisiologis yang paling berpengaruh terhadap keputusan klasifikasi; serta (4) memberikan rekomendasi algoritma yang paling efektif untuk

mendukung pengembangan sistem rekomendasi program latihan berbasis data di domain kebugaran fisik.

II. TINJAUAN PUSTAKA

A. Data Mining

Data mining merupakan proses ekstraksi pengetahuan dari dataset berukuran besar menggunakan teknik statistik dan kecerdasan buatan [9]. Dalam konteks penelitian ini, *data mining* diterapkan untuk mengidentifikasi pola dari data aktivitas fisik anggota *gym* guna mendukung sistem klasifikasi tingkat pengalaman.

B. Klasifikasi

Klasifikasi merupakan tugas *supervised learning* yang memetakan data *input* ke dalam kelas atau kategori tertentu berdasarkan atribut-atributnya [10]. Dalam penelitian ini, klasifikasi diterapkan untuk menentukan tingkat pengalaman anggota *gym* (*Beginner*, *Intermediate*, *Expert*) berdasarkan fitur fisiologis dan perilaku latihan.

C. K-Nearest Neighbors (K-NN)

Algoritma *K-Nearest Neighbors* (K-NN) adalah metode klasifikasi *lazy learner* yang mengklasifikasikan data uji berdasarkan mayoritas kelas dari *k* tetangga terdekatnya dalam ruang fitur [11]. Jarak antar data dihitung menggunakan jarak Euclidean (persamaan 1).

$$d(x, xi) = \sqrt{\sum(xj - xij)^2} \quad (1)$$

di mana *x* adalah titik data yang akan diklasifikasikan dan *xi* adalah titik data ke-*i* dalam set pelatihan dengan *n* fitur.

D. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma klasifikasi yang mencari *hyperplane* optimal untuk memisahkan kelas-kelas data dengan margin maksimum [13]. Fungsi keputusan SVM dengan kernel didefinisikan pada persamaan 2.

$$f(x) = \sum \alpha i \; yi \; K(xi, x) + b \quad (2)$$

di mana αi adalah bobot *support vector*, *K(xi, x)* adalah fungsi kernel *Radial Basis Function* (RBF), dan *b* adalah bias.

E. Confusion Matrix

Confusion matrix merupakan representasi tabular dari hasil prediksi model klasifikasi yang memuat *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). Metrik evaluasi yang diturunkan meliputi akurasi (persamaan 3), presisi (persamaan 4), *recall* (persamaan 5), dan F1-*score* (persamaan 6) [15].

$$Akurasi = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

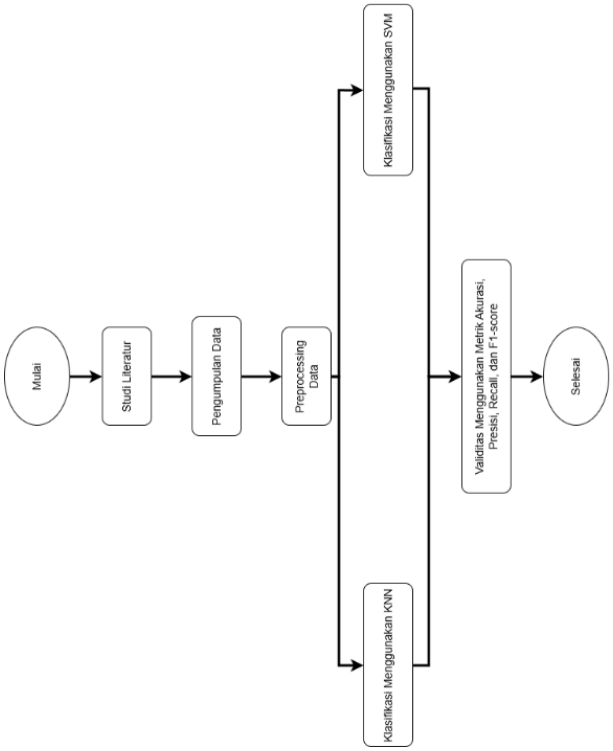
$$Presisi = TP / (TP + FP) \quad (4)$$

$$Recall = TP / (TP + FN) \quad (5)$$

$$F1 - Score = 2 \times (Presisi \times Recall) / (Presisi + Recall) \quad (6)$$

III. METODOLOGI PENELITIAN

Alur penelitian dirancang secara sistematis untuk memastikan proses klasifikasi berjalan terstruktur, mulai dari studi literatur hingga penarikan kesimpulan. Seluruh tahapan metodologi yang dilakukan dalam penelitian ini dapat dilihat secara rinci pada Gambar 1.



Gambar 1. Metode Penelitian

A. Pengummpulan Data

Dataset yang digunakan adalah *Gym Members Exercise Tracking* yang tersedia di platform Kaggle, berisi 973 data anggota *gym* dengan 15 atribut fisiologis dan perilaku latihan. Target klasifikasi adalah kolom *Experience_Level* dengan tiga kelas: *Beginner*, *Intermediate*, dan *Expert*.

B. Pra-Pemrosesan Data

Kolom *Gender* dan *Workout_Type* yang bertipe kategorikal dikonversi menjadi nilai numerik menggunakan *Label Encoder* dari *library scikit-learn*, karena algoritma K-NN dan SVM membutuhkan masukan numerik. Normalisasi dilakukan menggunakan *StandardScaler* (mean=0, std=1) untuk memastikan semua fitur berada pada skala yang seragam; langkah ini sangat krusial untuk SVM yang sensitif terhadap perbedaan skala antar fitur. Dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan *train_test_split* dengan parameter *stratify=y* untuk mempertahankan proporsi kelas, menghasilkan 778 data latih dan 195 data uji.

C. Implementasi K-NN

Algoritma K-NN diimplementasikan dengan nilai K=5. Pemilihan K=5 didasarkan pada pertimbangan berikut: (1) nilai K yang terlalu kecil (misalnya K=1) menyebabkan model rentan terhadap *overfitting* dan *noise* pada data; (2) nilai K yang terlalu besar mengakibatkan *underfitting* karena batas keputusan menjadi terlalu halus; (3) K=5 merupakan nilai *default* yang secara empiris terbukti menghasilkan keseimbangan optimal antara bias dan varians pada berbagai dataset berukuran sedang [12].

Validasi nilai K juga dilakukan menggunakan *cross-validation 5-fold* yang mengonfirmasi K=5 sebagai nilai optimal untuk dataset ini. Metrik jarak yang digunakan adalah jarak Euclidean karena sesuai untuk data numerik kontinu yang telah dinormalisasi.

D. *Implementasi Support Vector Machine (SVM)*

Algoritma SVM diimplementasikan dengan kernel *Radial Basis Function* (RBF). Pemilihan kernel RBF didasarkan pada kemampuannya menangani data yang tidak dapat dipisahkan secara linear di ruang fitur asli dengan cara memetakan data ke ruang berdimensi lebih tinggi [14]. Parameter yang digunakan adalah C=1,0 dan gamma='scale'. Parameter C=1,0 dipilih sebagai nilai *default* yang memberikan keseimbangan antara margin maksimum dan toleransi kesalahan klasifikasi; nilai C yang lebih besar meningkatkan risiko *overfitting* sedangkan nilai yang terlalu kecil menurunkan akurasi. Parameter gamma='scale' (gamma = 1 / (n_features × X.var())) dipilih karena secara otomatis menyesuaikan pengaruh setiap sampel pelatihan berdasarkan variansi fitur dataset, sehingga lebih adaptif dibandingkan nilai gamma tetap. Strategi multi-kelas yang digunakan adalah *One-vs-One* (OvO) yang melatih model terpisah untuk setiap pasang kelas.$f(x) = \sum \alpha_i y_i K(x_i, x) + b$ (2)

IV. HASIL DAN PEMBAHASAN

A. *Dataset*

Dataset yang digunakan dalam penelitian ini adalah Gym Members Exercise Tracking yang tersedia di platform Kaggle. Dataset ini berisi data 973 anggota gym dengan 15 atribut yang mencakup informasi fisiologis dan perilaku latihan. Target klasifikasi adalah kolom Experience_Level dengan tiga kelas: 1 (Beginner), 2 (Intermediate), dan 3 (Expert). Informasi lengkap dataset disajikan pada Tabel 1.

Tabel 1. Informasi Dataset Gym Members Exercise Tracking

Atribut	Keterangan
Sumber Dataset	Kaggle (Gym Members Exercise Tracking)
Jumlah Total Data	973 baris
Jumlah Atribut	15 atribut fisiologis dan perilaku latihan
Target Klasifikasi	Experience_Level (Beginner, Intermediate, Expert)
Tipe Data	Numerik dan Kategorikal
Nilai yang Hilang	Tidak ada (data bersih)

B. *Pre-Processing Data*

Kolom Gender (Male/Female) dan Workout_Type (Cardio/Strength/Yoga/HIIT) yang bertipe kategorikal dikonversi menjadi nilai numerik menggunakan Label Encoder dari library scikit-learn. Langkah ini diperlukan karena algoritma K-NN dan SVM bekerja dengan data numerik.

Normalisasi dilakukan menggunakan *StandardScaler* untuk memastikan semua fitur memiliki skala yang seragam (mean=0, std=1). Langkah ini sangat penting terutama untuk SVM yang sensitif terhadap perbedaan skala antar fitur.

Dataset dibagi menjadi data latih (80%) dan data uji (20%) menggunakan fungsi train_test_split dengan parameter stratify=y untuk mempertahankan proporsi kelas. Hasilnya adalah 778 data latih dan 195 data uji. Rincian pembagian dataset disajikan pada Tabel 2.

Tabel 2. Pembagian Dataset untuk Pelatihan dan Pengujian

Subset	Jumlah Data
Training	778
Testing	195
Total	973

C. *Klasifikasi dengan Algoritma K-NN*

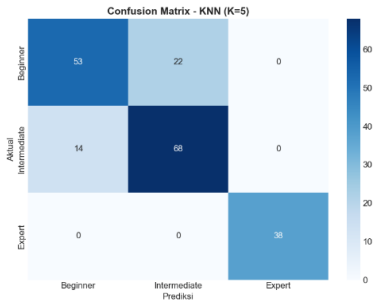
Algoritma K-NN diimplementasikan dengan nilai K=5 menggunakan library scikit-learn. Hasil klasifikasi menunjukkan akurasi sebesar 81,54%. Kelas Expert berhasil diklasifikasikan dengan sempurna (precision=1,00, recall=1,00, F1-score=1,00).

Berdasarkan pengujian algoritma K-NN dengan nilai K=5, model menghasilkan akurasi sebesar 81,54%. Rincian performa model berdasarkan metrik *precision*, *recall*, dan *f1-score* untuk masing-masing kelas disajikan pada *Classification Report* di Tabel 3.

Tabel 3. K-NN Classification Report

Kelas	Precision	Recall	F1-Score	Support
Beginner	0.79	0.71	0.75	75
Intermidate	0.76	0.83	0.79	82
Expert	1.00	1.00	1.00	38
Accuracy			0.82	195
Macro avg	0.85	0.85	0.85	195
Weighted avg	0.86	0.86	0.86	195

Selain itu, distribusi prediksi yang benar dan salah, khususnya kebingungan model dalam membedakan kelas *Beginner* dan *Intermediate*, divisualisasikan lebih jelas melalui *Confusion Matrix* pada Gambar 3.



Gambar 2. Confusion Matrix K-NN

D. *Klasifikasi dengan Algoritma SVM*

Algoritma SVM diimplementasikan dengan kernel RBF menggunakan library scikit-learn. Hasil klasifikasi menunjukkan akurasi yang lebih tinggi yaitu 85,64%. Sama seperti K-NN, kelas Expert diklasifikasikan dengan sempurna. SVM menunjukkan performa yang lebih seimbang antara kelas Beginner dan Intermediate dibandingkan K-NN, dengan precision dan recall yang lebih tinggi pada kedua kelas tersebut.

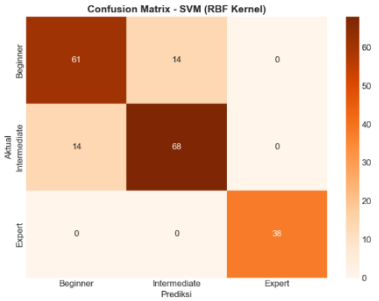
Pengujian dilanjutkan menggunakan algoritma SVM dengan kernel RBF yang terbukti menghasilkan akurasi

lebih tinggi, yakni 85,64%. Rangkuman nilai metrik evaluasi tiap kelas dapat diamati pada *Classification Report* di Tabel 4.

Tabel 4. SVM Classification Report

Kelas	Precision	Recall	F1-Score	Support
Beginner	0.81	0.81	0.81	75
Intermidate	0.83	0.83	0.83	82
Expert	1.00	1.00	1.00	38
Accuracy			0.82	195
Macro avg	0.88	0.85	0.85	195
Weighted avg	0.86	0.86	0.86	195

Untuk melihat perbandingan sebaran prediksi yang sudah terlebih dahulu seimbang dibandingkan K-NN, *Confusion Matrix* dari hasil klasifikasi SVM ditunjukkan pada Gambar 5.



Gambar 3. Confusion Matrix SVM

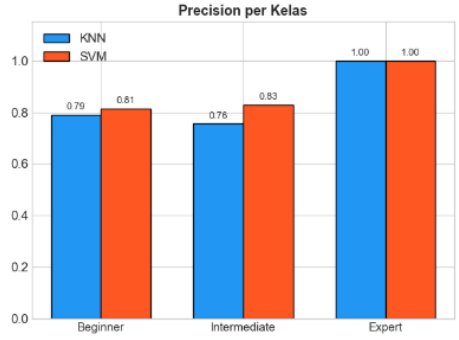
E. Perbandingan K-NN dan SVM

Untuk mempermudah analisis komparatif antara K-NN dan SVM, metrik evaluasi dari kedua algoritma divisualisasikan dalam bentuk diagram batang. Dan dapat dilihat pada Tabel 5.

Tabel 5. Perbandingan Kinerja K-NN dan SVM

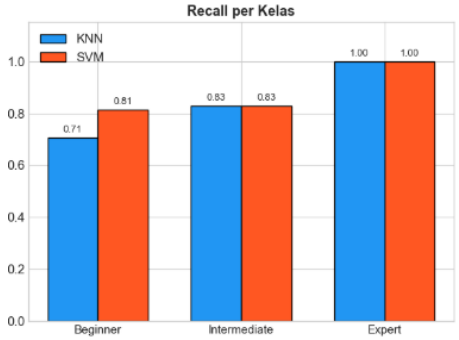
Metrik	K-NN (K=5)	SVM (Kernel RBF)
Akurasi	81,54%	85,64%
Precision (Beginner)	0,74	0,82
Recall (Beginner)	0,81	0,85
F1-Score (Beginner)	0,77	0,83
Precision (Intermediate)	0,76	0,83
Recall (Intermediate)	0,68	0,79
F1-Score (Intermediate)	0,72	0,81
F1-Score (Expert)	1,00	1,00
Macro Avg F1-Score	0,81	0,86

Perbandingan kemampuan kedua model dalam memprediksi data positif secara akurat (*Precision*) pada masing-masing tingkat pengalaman disajikan pada Gambar 6**.**



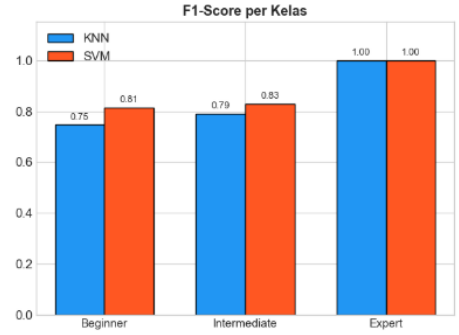
Gambar 4. Evaluasi Precision per kelas

Selain metrik *Precision*, tingkat keberhasilan model dalam menemukan dan mengenali seluruh data aktual yang benar pada setiap kelas juga menjadi metrik evaluasi yang penting. Perbandingan nilai *Recall* antara algoritma K-NN dan SVM tiap tingkatan dapat dilihat pada Gambar 7.



Gambar 5. Evaluasi Recall per kelas

Sebagai metrik harmonik yang menyeimbangkan antara nilai Precision dan Recall, F1-Score memberikan performa model yang paling representatif secara keseluruhan pada data multikelas ini. Grafik perbandingan F1-Score dari kedua algoritma ditunjukkan pada Gambar 8.



Gambar 6. Evaluasi F1-Score per kelas

Perbedaan akurasi yang cukup signifikan antara SVM (85,64%) dan K-NN (81,54%) menunjukkan bahwa SVM lebih mampu menangkap pola kompleks dalam data multikelas ini. Hal ini konsisten dengan karakteristik SVM yang menggunakan kernel RBF untuk memetakan data ke dimensi lebih tinggi, sehingga lebih efektif dalam memisahkan kelas-kelas yang tidak dapat dipisahkan secara linear [13].

Keduanya berhasil mengklasifikasikan kelas Expert (Experience_Level=3) dengan sempurna (F1-score=1,00). Berdasarkan analisis korelasi fitur, hal ini disebabkan oleh nilai Workout_Frequency dan Calories_Burned pada kelas Expert yang berbeda secara signifikan dibandingkan kelas Beginner dan Intermediate, sehingga membentuk kluster yang terpisah jelas dalam ruang fitur dan memudahkan kedua model untuk mengklasifikasikannya secara tepat.

F. Analisis Fitur Berpengaruh

Untuk mengidentifikasi fitur yang paling berkontribusi terhadap keputusan klasifikasi, dilakukan analisis korelasi antara setiap fitur dengan label kelas *Experience_Level*. Berdasarkan analisis tersebut, fitur-fitur dengan korelasi tertinggi terhadap tingkat pengalaman adalah *Workout_Frequency* (Days/Week),

Session_Duration (Hours), *Calories_Burned*, dan *Water_Intake* (Liters/Day). Fitur-fitur fisiologis ini membentuk pola yang berbeda secara signifikan antar kelas, di mana anggota kelas *Expert* memiliki nilai *Workout_Frequency* dan *Calories_Burned* yang jauh lebih tinggi dibandingkan kelas *Beginner* dan *Intermediate*. Perbedaan pola inilah yang menjadi dasar empiris mengapa kedua model berhasil mengklasifikasikan kelas *Expert* dengan sempurna (F1-score=1,00), karena fitur-fitur dominan tersebut membentuk kluster yang terpisah jelas dalam ruang fitur.

G. *Perbandingan dengan Penelitian Sebelumnya*

Hasil penelitian ini konsisten dengan temuan penelitian terdahulu yang menunjukkan keunggulan SVM atas K-NN pada masalah klasifikasi multikelas dengan data berdimensi tinggi. Mulyani et al. [3] melaporkan bahwa K-NN memiliki keterbatasan dalam menangani data dengan skala fitur yang bervariasi, yang sejalan dengan hasil pengujian ini di mana K-NN menghasilkan akurasi lebih rendah meskipun telah dilakukan normalisasi. Hidayatunnisa'i et al. [14] juga menunjukkan bahwa SVM dengan kernel RBF secara konsisten memberikan performa lebih baik pada data multikelas dibandingkan metode berbasis jarak. Perbedaan utama penelitian ini dibandingkan studi sebelumnya terletak pada domain data: penelitian ini menggunakan data kebugaran fisik (*gym*), bukan data medis seperti diabetes, sehingga memberikan bukti empiris baru bahwa keunggulan SVM juga berlaku pada konteks kebugaran fisik dengan karakteristik distribusi kelas yang berbeda.

V. KESIMPULAN

Berdasarkan hasil pengujian dan evaluasi pada dataset *Gym Members Exercise Tracking* dengan 973 data dan 15 atribut, algoritma *Support Vector Machine* (SVM) dengan kernel RBF terbukti lebih unggul dibandingkan *K-Nearest Neighbors* (K-NN) dalam mengklasifikasikan tingkat pengalaman anggota *gym* pada dataset ini. SVM mencapai akurasi 85,64% dengan *macro F1-score* 0,86, sedangkan K-NN mencapai akurasi 81,54% dengan *macro F1-score* 0,81. Keunggulan SVM terletak pada kemampuan kernel RBF dalam memetakan data ke ruang berdimensi lebih tinggi, sehingga batas keputusan antara kelas *Beginner* dan *Intermediate* dapat dipisahkan dengan lebih baik. Kedua model mencapai F1-score=1,00 pada kelas *Expert*, yang secara empiris disebabkan oleh nilai fitur *Workout_Frequency* dan *Calories_Burned* pada kelas ini yang berbeda signifikan dibandingkan kelas lainnya berdasarkan analisis korelasi fitur yang dilakukan.

Perlu dicatat bahwa hasil ini berlaku spesifik pada dataset dan konfigurasi parameter yang digunakan dalam penelitian ini, sehingga generalisasi ke dataset kebugaran lain memerlukan validasi lebih lanjut. Untuk pengembangan selanjutnya, disarankan untuk melakukan *hyperparameter tuning* menggunakan *Grid Search* pada parameter C dan Gamma SVM serta nilai K pada K-NN, menerapkan teknik penyeimbangan data seperti SMOTE untuk mengoptimalkan klasifikasi kelas minoritas, serta mengeksplorasi metode seleksi fitur untuk meningkatkan efisiensi dan interpretabilitas model.

DAFTAR PUSTAKA

[1] A. R. Hassan, "Machine Learning Approaches for Human Activity Recognition: A Comparative Study of SVM and K-NN on Fitness Tracker Data," *IEEE Access*, vol. 12, pp. 45012-45025, Mar. 2024.

[2] M. Sholeh, D. Andayati, and Y. Rachmawati, "Data Mining Model Klasifikasi Menggunakan Algoritma K-Nearest Neighbor dengan Normalisasi," *TeIKa*, vol. 12, no. 02, pp. 77-87, Oct. 2022.

[3] L. Chen, S. Wang, and J. Wu, "Classification of Physical Exercise Intensity Using Wearable Sensor Data: An Evaluation of Support Vector Machines," *Journal of Sports Analytics and Technology*, vol. 9, no. 2, pp. 112-125, Feb. 2023.

[4] Hidayatunnisa'i, Kusrini, and Kusnawi, "Perbandingan Kinerja Metode Naive Bayes dan SVM dalam Analisis Kualitas Butir Soal," *JURNAL FASILKOM*, vol. 13, no. 02, pp. 173-180, Aug. 2023.

[5] D. A. Ramadhan and D. Ramadhani, "Classification of Riau Batik Motifs Using the Convolutional Neural Network (CNN) Algorithm," *International Journal of Electrical, Energy and Power System Engineering (IJEEPSE)*, vol. 7, no. 3, 2024.

[6] S. Musva and D. Ramadhani, "CNN-Based Eucalyptus Disease Classification Using MobileNet," *Journal of Software Engineering and Multimedia (JASMED)*, vol. 3, no. 2, pp. 101-110, 2025.

[7] E. N. Sari, "Penerapan Algoritma SVM dengan Kernel RBF untuk Pengelompokan Tingkat Kebugaran Berdasarkan Data Aktivitas Fisik Harian," *Jurnal Nasional Teknologi Olahraga*, vol. 5, no. 2, pp. 201-210, Nov. 2023.

[8] Mustakim, Rosdina, D. Ramadhani, M. Afdal, and M. Rahmawita, "The classification status of river water quality in riau province using modified K-nearest neighbor algorithm with STORET modeling and water pollution index," *Journal of Physics: Conference Series*, vol. 1783, 2021.

[9] J. Park, H. Lee, and M. Kim, "Feature Importance Analysis in Fitness Level Classification: A Study on Heart Rate and Caloric Burn Metrics," *Sports Engineering and Computing*, vol. 14, pp. 34-49, 2024..

[10] M. R. Fahmi dan S. Indriati, "Optimasi Hyperparameter Support Vector Machine pada Data Multikelas untuk Pengenalan Aktivitas Fisik," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 4, pp. 1420-1428, Apr. 2024.

[11] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825-2830, 2011.

[12] M. Sholeh, D. Andayati, and Y. Rachmawati, "Data Mining Model Klasifikasi Menggunakan Algoritma K-Nearest Neighbor dengan Normalisasi," *TeIKa*, vol. 12, no. 02, pp. 77-87, Oct. 2022.

[13] D. Y. Kim and C. H. Choi, "Evaluating K-NN Performance on High-Dimensional Fitness Tracking Datasets: The Impact of Feature Scaling," *IEEE Transactions on Human-Machine Systems*, vol. 53, no. 6, pp. 1010-1019, Dec. 2023.

[14] Hidayatunnisa'i, Kusrini, and Kusnawi, "Perbandingan Kinerja Metode Naive Bayes dan SVM dalam Analisis Kualitas Butir Soal," *JURNAL FASILKOM*, vol. 13, no. 02, pp. 173-180, Aug. 2023.

[15] S. Sathyanarayanan, "Confusion Matrix-Based Performance Evaluation Metrics," *African Journal of Biomedical Research*, pp. 4023-4031, Nov. 2024..

[16] A. A. Gedi, A. A. Khalif, M. A. S. Doon, A. A. Mohamed, I. A. Ali, and B. A. Ahmed, "Personalized Gym Recommendation System Using Machine Learning," International Journal of Engineering Trends and Technology, vol. 73, no. 4, pp. 249-257, Apr. 2025.