

## **Efisiensi Pembakaran Kalori Menggunakan Algoritma Decision Tree pada Data Aktivitas Fisik**

**Ariska Nur Anggraini<sup>1\*</sup>, Endah Kurnia Fitri<sup>2</sup>**

<sup>1,2</sup>Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa, Indonesia

Dikirimkan: 06-04-2026  
Diterbitkan: 13-04-2026

**Keywords:**  
Decision Tree;  
Classification;  
Machine Learning;  
Calorie Burning.

**E-mail Penulis korespondensi:**  
ariskanur112@gmail.com

**Abstrak.** Penelitian ini bertujuan untuk mengklasifikasikan efisiensi pembakaran kalori berdasarkan data aktivitas fisik menggunakan algoritma Decision Tree. Dataset yang digunakan mencakup berbagai variabel aktivitas seperti durasi, intensitas, dan jenis aktivitas yang memengaruhi jumlah kalori yang terbakar. Proses klasifikasi dibagi menjadi tiga kategori utama, yaitu Low Efficiency, Moderate, dan High Efficiency. Model Decision Tree dibangun dan dievaluasi menggunakan metrik kinerja berupa akurasi, presisi, recall, dan F1-score. Hasil pengujian menunjukkan bahwa model memiliki tingkat akurasi sebesar 87,94%, dengan nilai presisi sebesar 89,43%, recall sebesar 87,94%, dan F1-score sebesar 88,66%. Meskipun secara keseluruhan performa model tergolong baik, analisis lebih lanjut menunjukkan adanya ketidakseimbangan kelas (class imbalance), di mana kategori Low Efficiency mendominasi data. Hal ini menyebabkan performa model pada kelas minoritas seperti High Efficiency dan Moderate relatif rendah, ditunjukkan oleh nilai presisi dan recall yang kecil pada kedua kelas tersebut. Dengan demikian, penelitian ini menunjukkan bahwa algoritma Decision Tree cukup efektif dalam mengklasifikasikan efisiensi pembakaran kalori, namun diperlukan penanganan lebih lanjut terhadap ketidakseimbangan data untuk meningkatkan performa klasifikasi pada seluruh kategori.

**Abstract.** This study aims to classify calorie-burning efficiency based on physical activity data using the Decision Tree algorithm. The dataset includes various activity-related variables such as duration, intensity, and type of activity that influence the number of calories burned. The classification is divided into three main categories: Low Efficiency, Moderate, and High Efficiency. The Decision Tree model was developed and evaluated using performance metrics including accuracy, precision, recall, and F1-score. The experimental results show that the model achieved an accuracy of 87.94%, with a precision of 89.43%, recall of 87.94%, and an F1-score of 88.66%. Although the overall performance of the model is considered good, further analysis reveals the presence of class imbalance, where the Low Efficiency category dominates the dataset. This imbalance leads to relatively poor performance in minority classes such as High Efficiency and Moderate, as indicated by low precision and recall values for these categories. Therefore, this study demonstrates that the Decision Tree algorithm is effective for classifying calorie-burning efficiency. However, further improvements are needed to address class imbalance in order to enhance classification performance across all categories.

### **1. Pendahuluan**

Aktivitas fisik merupakan salah satu faktor penting dalam menjaga kesehatan tubuh dan meningkatkan kualitas hidup [1]. Salah satu indikator yang sering digunakan untuk mengukur efektivitas aktivitas fisik adalah jumlah

kalori yang terbakar [2]. Efisiensi pembakaran kalori menjadi perhatian penting, terutama dalam konteks pengelolaan berat badan dan peningkatan kebugaran [3]. Dengan berkembangnya teknologi, data aktivitas fisik kini dapat dikumpulkan secara lebih mudah melalui berbagai perangkat seperti *wearable devices* dan aplikasi kebugaran.

Seiring dengan meningkatnya ketersediaan data aktivitas fisik, muncul peluang untuk memanfaatkan teknik *machine learning* dalam menganalisis dan mengklasifikasikan efisiensi pembakaran kalori. Pendekatan *machine learning* ini memungkinkan identifikasi pola-pola tertentu yang tidak mudah dikenali secara manual [4]. Salah satu metode yang sering digunakan dalam klasifikasi adalah algoritma Decision Tree, yang dikenal karena kemampuannya dalam menghasilkan model yang mudah dipahami dan diinterpretasikan [5].

Algoritma Decision Tree bekerja dengan membangun struktur pohon keputusan berdasarkan atribut-atribut yang ada dalam dataset [6]. Setiap node pada pohon merepresentasikan suatu kondisi atau pengujian terhadap atribut tertentu, sementara cabang menunjukkan hasil dari pengujian tersebut [7]. Metode ini memiliki keunggulan dalam menangani data numerik maupun kategorikal, serta mampu memberikan visualisasi yang intuitif terhadap proses pengambilan keputusan [8].

Dalam konteks klasifikasi efisiensi pembakaran kalori, penggunaan Decision Tree diharapkan dapat membantu dalam mengelompokkan data aktivitas fisik ke dalam kategori tertentu, seperti efisiensi rendah (*Low Efficiency*), sedang (*Moderate*), dan tinggi (*High Efficiency*). Klasifikasi ini dapat memberikan informasi yang bermanfaat bagi individu maupun praktisi kesehatan dalam merancang program latihan yang lebih efektif dan sesuai dengan kebutuhan.

Namun, dalam penerapan algoritma klasifikasi, sering kali ditemukan permasalahan seperti ketidakseimbangan data (*class imbalance*), di mana jumlah data pada setiap kelas tidak seimbang [9]. Kondisi ini dapat mempengaruhi kinerja model, terutama dalam memprediksi kelas minoritas. Akibatnya, model cenderung lebih akurat dalam memprediksi kelas mayoritas, namun kurang optimal pada kelas lainnya.

Selain itu, evaluasi kinerja model menjadi aspek penting dalam memastikan bahwa model yang dibangun memiliki kemampuan generalisasi yang baik. Metrik seperti akurasi, presisi, recall, dan F1-score digunakan untuk mengukur performa model secara menyeluruh [10]. Penggunaan beberapa metrik ini penting untuk memberikan gambaran yang lebih lengkap, terutama ketika menghadapi dataset yang tidak seimbang [11].

Penelitian ini bertujuan untuk menerapkan algoritma Decision Tree dalam mengklasifikasikan efisiensi pembakaran kalori berdasarkan data aktivitas fisik. Selain itu, penelitian ini juga menganalisis performa model yang dihasilkan serta mengidentifikasi tantangan yang muncul selama proses klasifikasi, khususnya yang berkaitan dengan distribusi data yang tidak merata.

Diharapkan hasil dari penelitian ini dapat memberikan kontribusi dalam pengembangan sistem berbasis *machine learning* untuk analisis aktivitas fisik. Selain itu, penelitian ini juga dapat menjadi referensi bagi penelitian selanjutnya dalam meningkatkan akurasi dan keandalan model klasifikasi, terutama dalam menangani permasalahan ketidakseimbangan data.

## 2. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk mengklasifikasikan efisiensi pembakaran kalori berdasarkan data aktivitas fisik menggunakan algoritma Decision Tree. Tahapan penelitian dilakukan secara sistematis mulai dari pengumpulan data hingga evaluasi model.

### 2.1. Alur Penelitian

Secara umum, alur penelitian ini meliputi beberapa tahapan utama yang dilakukan secara berurutan, yaitu pengumpulan data, pra-pemrosesan data, pembagian data, pembangunan model, evaluasi model, hingga analisis hasil. Tahap pengumpulan data dilakukan untuk memperoleh dataset aktivitas fisik yang relevan dan representatif [12]. Selanjutnya, data tersebut melalui tahap pra-pemrosesan untuk memastikan kualitas data, seperti pembersihan data, penanganan nilai yang hilang, serta transformasi data agar siap digunakan dalam proses pemodelan [13]. Setelah itu, data dibagi menjadi data latih dan data uji untuk keperluan pelatihan dan pengujian model [14]. Tahap pembangunan model dilakukan dengan menerapkan algoritma *Decision Tree* untuk membentuk struktur pohon keputusan yang mampu mengklasifikasikan data secara efektif. Model yang telah dibangun kemudian dievaluasi menggunakan berbagai metrik kinerja seperti akurasi, presisi, *recall*, dan *F1-score* guna mengetahui tingkat performanya [15]. Tahap terakhir adalah analisis hasil, yang bertujuan untuk menginterpretasikan performa model serta mengidentifikasi kelebihan dan kekurangannya, termasuk dalam

menghadapi ketidakseimbangan data [16]. Alur penelitian ini dirancang secara sistematis agar setiap tahapan saling terintegrasi dan menghasilkan model klasifikasi yang optimal serta dapat diandalkan.



Gambar 1. Alur dan Desain Penelitian

### 2.2. Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa data aktivitas fisik yang mencakup beberapa atribut seperti durasi aktivitas, intensitas, jenis aktivitas, serta jumlah kalori yang terbakar. Dataset diperoleh dari sumber sekunder (*kaggle*) yang relevan atau hasil pengumpulan melalui perangkat pelacak aktivitas (*fitness tracker*). Data kemudian disusun dalam bentuk dataset terstruktur untuk keperluan analisis.

### 2.3. Pra-pemrosesan Data

Tahap ini dilakukan untuk memastikan kualitas data sebelum digunakan dalam proses pemodelan. Proses yang dilakukan meliputi pembersihan data (*data cleaning*) untuk mengatasi nilai yang hilang (*missing values*), penghapusan data duplikat, serta transformasi data jika diperlukan [17]. Selain itu, dilakukan juga proses kategorisasi efisiensi pembakaran kalori menjadi tiga kelas, yaitu *Low Efficiency*, *Moderate*, dan *High Efficiency*.

### 2.4. Pembagian Data

Dataset yang telah diproses kemudian dibagi menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*). Pembagian ini bertujuan untuk melatih model serta menguji performanya terhadap data yang belum pernah dilihat sebelumnya [18]. Umumnya, proporsi pembagian yang digunakan adalah 80% untuk data latih dan 20% untuk data uji.

### 2.5. Pemodelan Decision Tree

Model klasifikasi dibangun menggunakan algoritma Decision Tree. Algoritma ini bekerja dengan membentuk struktur pohon berdasarkan atribut yang paling berpengaruh terhadap target klasifikasi [19]. Proses pelatihan dilakukan menggunakan data latih hingga terbentuk model yang mampu mengklasifikasikan data ke dalam kategori efisiensi yang telah ditentukan.

### 2.6. Evaluasi Model (*Model Evaluation*)

Evaluasi model dilakukan menggunakan data uji untuk mengukur kinerja model yang telah dibangun. Metrik evaluasi yang digunakan meliputi akurasi, *presisi*, *recall*, dan *F1-score*. Selain itu, digunakan juga *classification report* untuk melihat performa model pada masing-masing kelas, terutama dalam mengidentifikasi ketidakseimbangan data (*class imbalance*) [20].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

2.6. Analisis Hasil

Hasil evaluasi model dianalisis untuk mengetahui kelebihan dan kekurangan dari algoritma yang digunakan [21]. Dalam penelitian ini, perhatian khusus diberikan pada performa model terhadap kelas minoritas yang cenderung memiliki nilai evaluasi yang lebih rendah. Analisis ini digunakan sebagai dasar untuk memberikan rekomendasi perbaikan pada penelitian selanjutnya.

3. Hasil dan Pembahasan

Pada penelitian ini, model klasifikasi efisiensi pembakaran kalori dibangun menggunakan algoritma Decision Tree berdasarkan data aktivitas fisik. Proses pelatihan dan pengujian model dilakukan menggunakan dataset yang telah melalui tahap pra-pemrosesan dan pembagian data. Hasil evaluasi model menunjukkan bahwa algoritma Decision Tree mampu menghasilkan performa yang cukup baik secara keseluruhan.

Tabel 1. Dataset Penelitian

Age	Steps/Day	Calories Burned	...	Sleep Hours	Workout/Week	...	Exercise Days	Efficiency Score	Kategori
51	7853	1500	...	6.42	3	...	1	0.603	Low Efficiency
60	4820	1500	...	6.82	6	...	3	0.958	Low Efficiency
59	4251	1500	...	6.99	2	...	1	0.987	Low Efficiency
39	6275	1500	...	6.65	1	...	1	0.711	Low Efficiency
22	6490	1500	...	5.80	2	...	6	0.551	Low Efficiency
53	8517	1500	...	6.42	3	...	2	0.754	Low Efficiency
23	8263	1500	...	5.11	2	...	1	0.492	Low Efficiency
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
29	6901	1500	...	6.11	7	...	1	0.426	Low Efficiency

Tabel 1 merupakan contoh data mentah (raw data) yang digunakan dalam penelitian. Dataset terdiri dari berbagai atribut fisiologis dan aktivitas fisik yang berpengaruh terhadap efisiensi pembakaran kalori. Kolom terakhir menunjukkan label kelas (*calorie efficiency*) yang menjadi target dalam proses klasifikasi menggunakan algoritma *Decision Tree*.

Tabel 2. Classification Report

Kelas	Precision	Recall	F1-Score	Support
High Efficiency	0.03	0.04	0.03	5,379
Low Efficiency	0.95	0.93	0.94	187,649
Moderate	0.18	0.23	0.21	6,972
Accuracy			0.88	200,000
Macro Avg	0.39	0.40	0.39	200,000
Weighted Avg	0.89	0.88	0.89	200,000

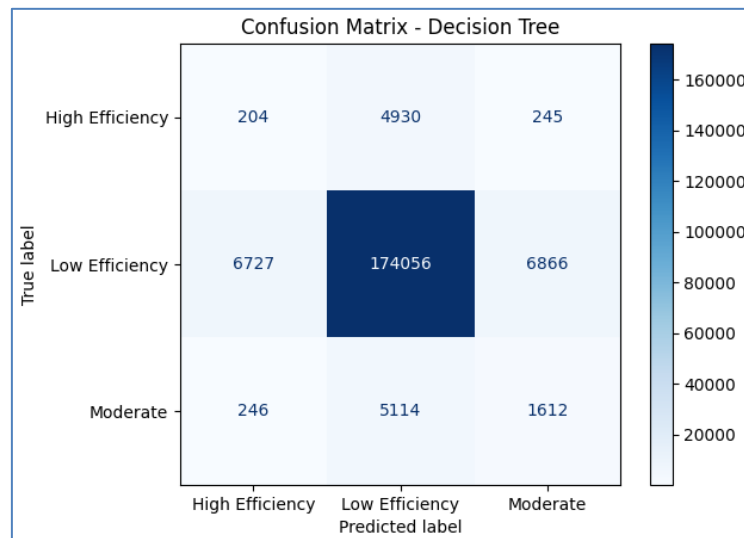
Berdasarkan hasil pengujian, model memperoleh nilai akurasi sebesar 87,94%, yang menunjukkan bahwa sebagian besar data berhasil diklasifikasikan dengan benar. Selain itu, nilai presisi sebesar 89,43% menunjukkan bahwa model memiliki kemampuan yang baik dalam meminimalkan kesalahan prediksi positif. Nilai recall sebesar 87,94% mengindikasikan bahwa model cukup mampu dalam mengenali data dari masing-masing kelas, sedangkan nilai F1-score sebesar 88,66% menunjukkan keseimbangan yang baik antara presisi dan recall.

Namun, jika dilihat lebih rinci melalui *classification report*, terdapat perbedaan performa yang signifikan antar kelas. Kelas *Low Efficiency* memiliki nilai presisi dan recall yang sangat tinggi, yaitu masing-masing sebesar 0,95 dan 0,93, yang menunjukkan bahwa model sangat baik dalam mengenali kelas mayoritas ini. Hal ini disebabkan oleh dominasi jumlah data pada kelas tersebut dibandingkan dengan kelas lainnya.

Sebaliknya, performa model pada kelas *Moderate* dan *High Efficiency* tergolong rendah. Kelas *Moderate* hanya memiliki nilai presisi sebesar 0,18 dan recall sebesar 0,23, sedangkan kelas *High Efficiency* memiliki nilai presisi sebesar 0,03 dan recall sebesar 0,04. Rendahnya nilai ini menunjukkan bahwa model kesulitan dalam mengidentifikasi kelas minoritas, sehingga banyak data dari kelas tersebut yang salah diklasifikasikan.

Permasalahan ini disebabkan oleh ketidakseimbangan data (*class imbalance*), di mana jumlah data pada kelas *Low Efficiency* jauh lebih besar dibandingkan kelas lainnya. Kondisi ini menyebabkan model cenderung “bias” terhadap kelas mayoritas dan mengabaikan pola pada kelas minoritas. Hal ini juga diperkuat dengan munculnya peringatan (*warning*) terkait metrik presisi yang tidak terdefinisi pada beberapa label karena tidak adanya prediksi pada kelas tertentu.

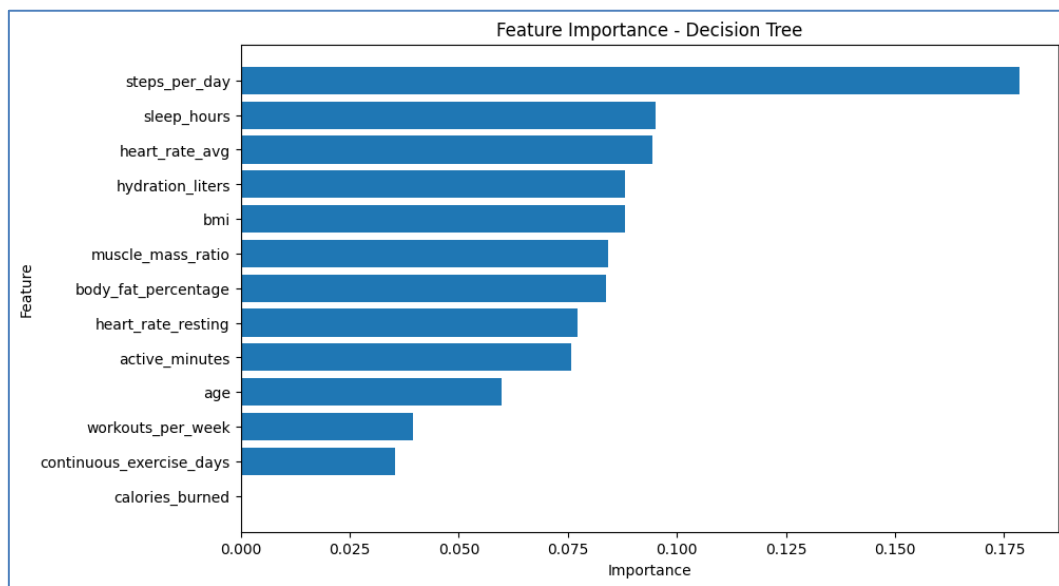
Nilai rata-rata makro (*macro average*) yang berada di sekitar 0,39 menunjukkan bahwa performa model secara keseluruhan sebenarnya kurang merata di semua kelas. Sementara itu, nilai rata-rata tertimbang (*weighted average*) yang tinggi, yaitu sekitar 0,89, lebih dipengaruhi oleh dominasi kelas mayoritas sehingga tidak sepenuhnya mencerminkan performa model pada seluruh kelas.



Gambar 2. Confusion Matrix

Berdasarkan hasil *confusion matrix*, dapat dilihat bahwa model *Decision Tree* memiliki performa yang sangat baik dalam mengklasifikasikan kelas *Low Efficiency*, ditunjukkan oleh jumlah prediksi benar yang sangat tinggi yaitu sebanyak 174.056 data. Namun, performa model pada kelas *High Efficiency* dan *Moderate* masih tergolong rendah. Hal ini terlihat dari jumlah prediksi benar pada kelas *High Efficiency* yang hanya sebesar 204 data, serta kelas *Moderate* sebesar 1.612 data. Selain itu, terdapat banyak kesalahan klasifikasi di mana data dari kelas *High Efficiency* dan *Moderate* cenderung diprediksi sebagai *Low Efficiency*, masing-masing sebanyak 4.930 dan 5.114 data. Kondisi ini menunjukkan bahwa model memiliki kecenderungan bias terhadap kelas mayoritas, yaitu *Low Efficiency*, yang jumlah datanya jauh lebih besar dibandingkan kelas lainnya. Akibatnya, model kurang mampu mengenali pola pada kelas minoritas. Fenomena ini sejalan dengan hasil *classification report* yang menunjukkan rendahnya nilai *precision* dan *recall* pada kelas *High Efficiency* dan *Moderate*. Oleh karena itu, dapat disimpulkan bahwa meskipun model memiliki akurasi tinggi, performa klasifikasi belum merata pada semua kelas. Untuk meningkatkan kinerja model, diperlukan pendekatan tambahan seperti penanganan ketidakseimbangan data menggunakan teknik *oversampling*, *undersampling*, atau penerapan algoritma lain yang lebih robust terhadap imbalanced data.

Berdasarkan hasil analisis *feature importance* pada model *Decision Tree*, dapat diketahui bahwa atribut yang paling berpengaruh dalam menentukan klasifikasi efisiensi pembakaran kalori adalah *steps\_per\_day*, dengan nilai kepentingan tertinggi dibandingkan fitur lainnya. Hal ini menunjukkan bahwa jumlah langkah per hari merupakan faktor utama dalam menentukan efisiensi pembakaran kalori. Selanjutnya, fitur seperti *sleep\_hours*, *heart\_rate\_avg*, *hydration\_liters*, dan *bmi* juga memiliki kontribusi yang cukup signifikan dalam proses klasifikasi. Fitur-fitur tersebut berkaitan langsung dengan kondisi fisiologis dan kebiasaan hidup yang memengaruhi metabolisme tubuh. Selain itu, atribut seperti *muscle\_mass\_ratio*, *body\_fat\_percentage*, dan *heart\_rate\_resting* juga memberikan pengaruh yang moderat terhadap model. Sementara itu, fitur seperti *active\_minutes*, *age*, *workouts\_per\_week*, dan *continuous\_exercise\_days* memiliki tingkat kontribusi yang lebih rendah. Menariknya, variabel *calories\_burned* justru memiliki pengaruh yang sangat kecil, yang kemungkinan disebabkan oleh nilai yang relatif konstan dalam dataset. Secara keseluruhan, hasil ini menunjukkan bahwa aktivitas harian dan kondisi fisik individu memiliki peran penting dalam menentukan efisiensi pembakaran kalori. Analisis ini juga membantu dalam memahami variabel mana yang paling relevan, sehingga dapat digunakan sebagai dasar dalam pengembangan model yang lebih optimal di masa mendatang.



Gambar 3. Feature Importance

Berdasarkan hasil tersebut, dapat disimpulkan bahwa algoritma *Decision Tree* cukup efektif dalam menangani klasifikasi secara umum, namun kurang optimal dalam menghadapi dataset yang tidak seimbang. Oleh karena itu, diperlukan pendekatan tambahan seperti teknik *resampling* (*oversampling* atau *undersampling*), penggunaan algoritma yang lebih *robust*, atau penyesuaian parameter model untuk meningkatkan performa pada kelas minoritas.

#### 4. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *Decision Tree* mampu digunakan untuk mengklasifikasikan efisiensi pembakaran kalori berdasarkan data aktivitas fisik dengan performa yang cukup baik secara keseluruhan. Hal ini ditunjukkan oleh nilai akurasi sebesar 87,94%, serta nilai presisi, recall, dan F1-score yang relatif tinggi. Model mampu mengidentifikasi pola dari data dan memberikan hasil klasifikasi yang akurat, khususnya pada kelas mayoritas. Namun demikian, hasil analisis menunjukkan adanya ketidakseimbangan data (*class imbalance*) yang signifikan, di mana kelas *Low Efficiency* mendominasi dataset. Kondisi ini menyebabkan model memiliki performa yang sangat baik pada kelas tersebut, tetapi kurang optimal dalam mengklasifikasikan kelas minoritas seperti *High Efficiency* dan *Moderate*. Hal ini terlihat dari rendahnya nilai *precision* dan *recall* pada kedua kelas tersebut serta dari hasil *confusion matrix* yang menunjukkan banyaknya kesalahan klasifikasi ke kelas mayoritas. Selain itu, analisis *feature importance* menunjukkan bahwa variabel seperti *steps\_per\_day*, *sleep\_hours*, dan *heart\_rate\_avg* memiliki pengaruh besar dalam menentukan efisiensi pembakaran kalori. Sebaliknya, beberapa variabel lain seperti *calories\_burned* memiliki kontribusi yang relatif kecil terhadap model. Dengan demikian, meskipun algoritma *Decision Tree* efektif dalam klasifikasi secara umum, diperlukan upaya lebih lanjut untuk meningkatkan performa model, terutama pada kelas minoritas. Beberapa pendekatan yang dapat dilakukan antara lain penerapan teknik penyeimbangan data seperti *oversampling* dan *undersampling*, serta penggunaan metode klasifikasi lain yang lebih *robust* terhadap data yang tidak seimbang. Penelitian ini diharapkan dapat menjadi dasar bagi pengembangan model yang lebih akurat dan adil dalam klasifikasi efisiensi pembakaran kalori di masa mendatang.

#### Daftar Rujukan

- [1] Lisa Khoirunnisa and Nur Muhammad Soleh, "PENGARUH POLA HIDUP SEHAT TERHADAP KESEHATAN FISIK DAN MENTAL," *J. Cent. Publ.*, vol. 2, no. 2, pp. 1686–1691, Apr. 2025, doi: 10.60145/jcp.v2i2.352.
- [2] W. H. Al Qindy, J. Y. Damaira, and S. H. L. Siregar, "Analisis langkah jalan bulanan dan kalori yang dihasilkan pada pemahaman konsep energi dan metabolisme individu," *Tak. J. Pendidikan, Sains, dan Hum.*, vol. 4, no. 2, pp. 247–256, Aug. 2025, doi: 10.56113/takuana.v4i2.143.
- [3] F. A. Tjakradidjaja, "Puasa Intermiten dan Transformasi Tubuh: Sains di Balik Penurunan Lemak, Pemeliharaan Otot, dan Optimalisasi Energi," *Asian J. Multidiscip. Res.*, vol. 1, no. 4, pp. 157–165, Nov. 2024, doi: 10.59613/db4jpd74.
- [4] J. J. Hidayat and S. Hasanudin, "Prediksi Volatilitas IHSG Dengan Hybrid Model GARCH–Random Forest Berbasis Machine Learning," *J. Manaj. Inform. Teknol.*, vol. 6, no. 1 SE-Articles, pp. 130–140, 2026, doi: 10.51903/mifortekh.v6i1.1134.
- [5] J. R. Quinlan, "Learning decision tree classifiers," *ACM Comput. Surv.*, vol. 28, no. 1, pp. 71–72, Mar. 1996, doi: 10.1145/234313.234346.

- [6] C. N. Syahputri and M. S. Hasibuan, "OPTIMASI KLASIFIKASI DECISION TREE DENGAN TEKNIK PRUNING UNTUK MENGURANGI OVERFITTING," *JSii (Jurnal Sist. Informasi)*, vol. 11, no. 2, pp. 87–96, Sep. 2024, doi: 10.30656/jsii.v11i2.9161.
- [7] A. Ermillian and K. Nugroho, "Perancangan Model Deteksi Potensi Siswa Putus Sekolah Menggunakan Metode Logistic Regression Dan Decision Tree," *J. Inform. J. Pengemb. IT*, vol. 9, no. 3, pp. 281–295, Dec. 2024, doi: 10.30591/jpit.v9i3.8007.
- [8] A. Riansah, O. Nurdian, and R. Herdiana, "PENERAPAN ALGORITMA RANDOM FOREST DAN DECISION TREE UNTUK MENINGKATKAN AKURASI KLASIFIKASI PENJUALAN PADA TOKO BANGUNAN," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 3, pp. 4242–4249, May 2025, doi: 10.36040/jati.v9i3.13622.
- [9] C. V. Angkoso, K. Asror, A. Kusumaningsih, and A. K. Nugroho, "Optimasi Algoritma Support Vector Machine Berbasis Kernel Radial Basis Function (RBF) Menggunakan Metode Particle Swarm Optimization Untuk Analisis Sentimen," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 12, no. 3, pp. 705–718, Jun. 2025, doi: 10.25126/jtiik.2025129317.
- [10] M. D. I. Amin, J. J. Hidayat, C. Setyowati, E. K. Fitri, A. N. Anggraini, and A. P. Werdana, "Implementasi Model LSTM Untuk Peramalan Curah Hujan Di Bekasi Dengan Pemanfaatan Data Cuaca BMKG," *J. Teknol. Inf. Digit.*, vol. 1, no. 2, pp. 90–99, [Online]. Available: <https://jurnal.ipdig.id/index.php/jtid/article/view/200>
- [11] A. M. Rifa'i, M. A. Burhanuddin, W. Hadikristanto, A. Saputra, D. Ariatmanto, and M. S. Mozamir, "Improving Blood Cell Subtype Classification Performance Using Data Augmentation with MobileNetV2," in *2025 International Conference on Smart Computing, IoT and Machine Learning (SIML)*, IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/SIML65326.2025.11081145.
- [12] A. M. Rifai, S. Raharjo, E. Utami, and D. Ariatmanto, "Analysis for diagnosis of pneumonia symptoms using chest X-ray based on MobileNetV2 models with image enhancement using white balance and contrast limited adaptive histogram equalization (CLAHE)," *Biomed. Signal Process. Control*, vol. 90, p. 105857, Apr. 2024, doi: 10.1016/j.bspc.2023.105857.
- [13] J. J. Hidayat, C. Setyowati, M. D. I. Amin, K. Bimasakti, and A. P. Werdana, "Deep Learning-based Sentiment Analysis of Public Comments on Military Education Using RoBERTa Algorithm and Rule-Based Hybrid Parameters," *J. Media Comput. Sci.*, vol. 4, no. 2 SE-Articles, Jul. 2025, doi: 10.37676/jmcs.v4i2.8769.
- [14] J. J. Hidayat, M. D. I. Amin, E. K. Fitri, A. N. Anggraini, A. P. Werdana, and C. Setyowati, "Implementasi Model EfficientNetB0 Pada Pembuatan Aplikasi Desktop Untuk Identifikasi Hama Tanaman Sawi Berbasis Deep Learning," *J. Teknol. Inf. Digit.*, vol. 1, no. 2, pp. 82–89, [Online]. Available: <https://jurnal.ipdig.id/index.php/jtid/article/view/195>
- [15] A. H. Anshor and A. T. Zy, "Implementasi Metode Decision Tree pada Sistem Prediksi Status Kualitas Produk Minuman A," *J. Ilm. Inform. Glob.*, vol. 15, no. 1, pp. 17–22, Apr. 2024, doi: 10.36982/jiig.v15i1.3778.
- [16] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, p. 27, Dec. 2019, doi: 10.1186/s40537-019-0192-5.
- [17] C. Setyowati, A. N. Anggraini, and E. K. Fitri, "Analisis Sentimen dan Karakteristik Linguistik Komentar Publik terhadap Kebijakan Militer Menggunakan Model RoBERTa," *J. Teknol. Inf. Digit.*, vol. 2, no. 1, pp. 39–46, [Online]. Available: <https://jurnal.ipdig.id/index.php/jtid/article/view/219>
- [18] A. Putra Argadinata, D. Abdul Fatah, and H. Sukri, "KLASIFIKASI KUALITAS BUAH APEL MENGGUNAKAN METODE RANDOM FOREST," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 2, pp. 2016–2022, Mar. 2025, doi: 10.36040/jati.v9i2.12854.
- [19] A. Reza, A. N. Al-kharits, M. Bustomi, N. Maulana, and T. A. Rahman, "Diabetes Prediction Based on Patient Medical Factors Using Decision Tree Algorithm," *J. Inf. Technol. Informatics Eng.*, vol. 1, no. 1, pp. 30–36, [Online]. Available: <https://journal.jci.co.id/jitie/article/view/123>
- [20] J. J. Hidayat, C. Setyowati, and A. P. Werdana, "Perancangan Sistem Prediksi Penyakit pada Tanaman Padi Berbasis Image Processing Menggunakan Algoritma VGG-16 Transfer Learning dan K-Means Segmentation," *J. Pract. Comput. Sci.*, vol. 5, no. 1, pp. 1–15, May 2025, doi: 10.37366/jpcs.v5i1.5759.
- [21] F. Putra, H. F. Tahiyat, R. M. Ihsan, R. Rahmadden, and L. Efrizoni, "Penerapan Algoritma K-Nearest Neighbor Menggunakan Wrapper Sebagai Preprocessing untuk Penentuan Keterangan Berat Badan Manusia," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 1, pp. 273–281, Jan. 2024, doi: 10.57152/malcom.v4i1.1085.