

# IMPLEMENTASI METODE MACHINE LEARNING UNTUK MENGKLASIFIKASI AKTIVITAS FISIK PADA REMAJA BERBASIS DATA KUESIONER

Oleh

Dini Siti Cahya Kartini\*<sup>1</sup>, Jajat<sup>2</sup>, Yati Ruhayati<sup>3</sup>, Kuston Sultoni<sup>4</sup>, Adang  
Suherman<sup>5</sup>, Imas Damayanti<sup>6</sup>, Nur Indri Rahayu<sup>7</sup>

<sup>1,2,3,4,5,6,7</sup> Program Studi Ilmu Keolahragaan Universitas Pendidikan Indonesia

Email: [dinisiti.21@upi.edu](mailto:dinisiti.21@upi.edu), [jajat\\_kurdul@upi.edu](mailto:jajat_kurdul@upi.edu), [yatiruhayati@upi.edu](mailto:yatiruhayati@upi.edu),  
[kuston.sultoni@upi.edu](mailto:kuston.sultoni@upi.edu), [adangsuherman@upi.edu](mailto:adangsuherman@upi.edu), [imas\\_d@upi.edu](mailto:imas_d@upi.edu),  
[nurindirahayu1910@upi.edu](mailto:nurindirahayu1910@upi.edu)

## Abstrak

Penggunaan *artificial intelligence* (AI) dalam berbagai bidang kehidupan termasuk aktivitas fisik dan olahraga menjadi salah satu yang sedang trending pada saat ini. Adapun tujuan penelitian ini adalah mengklasifikasi aktivitas fisik dengan metode *machine learning* berbasis data kuesioner berdasarkan waktu aktivitas dan *Metabolic Equivalent of Task* (MET). Subjek penelitian ini adalah 779 orang remaja usia 17-21 tahun ( $M \pm SD = 19,34 \pm 0,39$ ) yang berasal dari siswa SMA dan Mahasiswa di Jawa Barat. *International Physical Activity Questionnaire – Short Form* (IPAQ SF) digunakan untuk mengumpulkan data aktivitas fisik. Adapun algoritma *machine learning* yang digunakan yaitu *decision tree*. Hasil dari penelitian ini menunjukkan bahwa akurasi performa *decision tree* untuk mengklasifikasi aktivitas fisik berdasarkan variabel atribut kalkulasi MET lebih tinggi dibandingkan dengan atribut waktu aktivitas fisik (93% ; 86,67%). Algoritma *decision tree* memiliki akurasi tinggi dalam mengklasifikasi aktivitas fisik dengan atribut kalkulasi MET di setiap level. Analisis lebih lanjut dengan algoritma berbeda diperlukan untuk mengkaji performa terbaik.

**Kata kunci:** aktivitas fisik, *artificial intelligence*, MET, *decision tree*

## A. PENDAHULUAN

Perkembangan zaman dan teknologi saat ini berkembang demikian pesat, sehingga memberikan kemudahan dalam memperoleh akses (Hidayah et al., 2019) baik komunikasi, transportasi, kebutuhan pangan, dan lainnya dapat diperoleh dengan mudah hanya dengan mengandalkan *smartphone* (Hasan et al., 2019). Namun demikian, perkembangan teknologi juga memberikan dampak negatif, seperti halnya kebiasaan kecenderungan dengan penggunaan *smartphone* yang terjadi di kalangan anak-anak dan remaja, sehingga mengakibatkan terjadinya kebiasaan sedentary lifestyle (Gustra Jeki et al., 2022). Di sisi lain, perkembangan teknologi yang memberikan banyak kemudahan

---

hal ini juga dapat mengakibatkan penurunan dalam aktivitas fisik atau gerak (Zwingly C et al., 2019) dan berisiko terjadinya penyakit tidak menular (Marbun et al., 2021).

Sebagaimana hasil kajian dari World Health Organization (WHO) bahwa penyebab kematian terbanyak di dunia termasuk di Indonesia saat ini disebabkan oleh penyakit tidak menular seperti penyakit ginjal kronis, diabetes melitus, dan hipertensi (WHO, 2022). Adapun penyakit - penyakit itupun disebabkan faktor rendahnya aktivitas fisik dan tingginya perilaku sedentari. Sementara itu aktivitas fisik yang teratur dapat bermanfaat terhadap kesehatan biologis / fisik dan kesehatan psikologis untuk kesehatan jiwa pada semua jenjang usia (Rohmah et al., 2021; wilda welis, 2013), khususnya pembiasaan pada masa kanak-kanak, remaja menuju dewasa. Pada fase tersebut terjadi perubahan secara fisik, mental maupun sosial (Dumith et al., 2011). Adapun aktivitas fisik yang bermanfaat dan direkomendasikan untuk anak dan remaja usia 5-17 tahun, melakukan setidaknya rata-rata 60 menit per hari intensitas sedang hingga tinggi. Selain aktivitas aerobik intensitas tinggi, juga dianjurkan untuk melakukan aktivitas yang memperkuat otot dan tulang setidaknya 3 hari seminggu (WHO, 2020).

Salah satu permasalahan mengenai apakah aktivitas fisik itu termasuk ke dalam kategori aktivitas fisik rendah, sedang atau tinggi belum dipahami oleh setiap orang, meskipun banyak referensi mengenai jenis aktivitas disertai dengan prediksi klasifikasinya. Selain itu, untuk mengetahui level atau intensitas aktivitas fisik salah satu metode yang sering digunakan yaitu dengan menggunakan kuesioner. Kuesioner aktivitas fisik biasanya berisi pertanyaan-pertanyaan tentang jenis aktivitas sehari-hari, frekuensi, durasi, dan intensitas aktivitas fisik dan *metabolic equivalent of task* yang digunakan untuk menentukan level aktivitas fisik (Rondonuwu et al., 2020). Selain dengan kuesioner, klasifikasi aktivitas fisik juga telah banyak yang menggunakan accelerometer seperti Actigraph dan ActivPal yang memiliki tingkat objektivitas lebih tinggi.

Menurut beberapa ahli, kuesioner memiliki beberapa kelemahan, seperti jawaban sering tidak objektif, terutama bila pertanyaan kurang tajam sehingga memungkinkan responden berpura-pura atau ketidak lengkapan data (Alexander et al., 2020). Namun, Kuesioner juga memiliki beberapa kelebihan seperti menghemat waktu, menghemat akurat dalam mengukur aktivitas yang intens biaya, menghemat tenaga, mampu menjangkau responden yang cukup luas dapat membuat peringkat individu atau kelompok dalam aktivitas fisik (Besson et al., 2010; Sauro & Lewis, 2011).

IPAQ sendiri merupakan kuesioner untuk mengukur aktivitas fisik yang telah banyak digunakan di sejumlah negara termasuk Indonesia karena efisiensi dan kemudahannya. IPAQ menanyakan aktivitas yang telah dilakukan oleh responden dalam 7 (tujuh) hari ke belakang disertai dengan lama waktu dalam melakukannya. Jawaban dari responden selanjutnya akan dikalkulasi dengan rumus, sehingga diperoleh prediksi MET untuk selanjutnya diperoleh klasifikasi dari aktivitas fisik. Namun

---

demikian, dikarenakan yang ditanyakan adalah aktivitas dalam 7 (tujuh) hari ke belakang, terkadang membuat responden terkadang mengalami kesulitan untuk memberikan jawaban sehingga jawabannya tidak akurat. Jawaban yang tidak sesuai ini pada akhirnya memiliki kecenderungan juga salah (Chastin et al., 2014) dalam mengklasifikasi aktivitas fisik. Dengan demikian perlu adanya kajian mengenai klasifikasi aktivitas fisik berbasis kuesioner secara lebih mendalam.

Pada saat ini banyak penelitian yang mengkaji aktivitas fisik, baik itu memprediksi maupun mengklasifikasi dengan memanfaatkan teknologi *artificial Intelligence* (AI) (Miric et al., 2023). Dari AI selanjutnya berkembang metode *machine learning* dengan berbagai algoritma yang dapat digunakan untuk mengklasifikasi maupun memprediksi aktivitas fisik. Metode *machine learning* dapat digunakan untuk memprediksi dan mengklasifikasi aktivitas fisik berdasarkan data yang diambil dari berbagai instrument. Tahapannya dimulai dari pengumpulan data, anotasi data, pra pemrosesan data, pemilihan fitur, pembagian data, penentuan algoritma yang akan digunakan, pelatihan model, dan evaluasi performa model atau testing data (Hüllermeier, 2005). Metode klasifikasi pada *ML* yang sering digunakan pada metode *machine learning* adalah klasifikasi and Regression Trees (CART), decision tree, random forest, naïve bayes, *support vector machines* (SVM), dan lain-lain yang secara bertahap meningkatkan akurasinya (Sihombing & Yuliati, 2021).

Saat ini di Indonesia juga telah banyak penelitian yang memanfaatkan *machine learning* dalam mengklasifikasi dan memprediksi berbagai bidang keilmuan. Namun demikian, penelitian yang mengklasifikasi aktivitas fisik berbasis kuesioner IPAQ dengan variabel atribut waktu dan MET relatif masih jarang. Oleh karena itu penelitian ini bertujuan untuk menganalisis klasifikasi aktivitas fisik berbasis data kuesioner pada usia remaja dengan atribut waktu aktivitas yang dilaporkan dan kalkulasi MET.

## **B. METODE PENELITIAN**

### **Study Design**

Sumber data penelitian ini diperoleh dari data set payung penelitian dosen dan mahasiswa program studi Ilmu Keolahragaan (IKOR) Fakultas Pendidikan Olahraga dan Kesehatan (FPOK) Universitas Pendidikan Indonesia (UPI) yang tergabung dalam kelompok bidang keilmuan aktivitas fisik. Pengambilan data ini dilaksanakan pada sejak tahun 2020 sampai 2023. Model klasifikasi pada penelitian ini, dilakukan melalui beberapa proses: (1) Physical activity data set; (2) METs data; (3) preprocessing data; (4) Model Klasifikasi (5) Steps Data

### **Physical Activity Data Set**

Data yang digunakan dalam penelitian bersumber dari beberapa penelitian kelompok bidang keilmuan aktivitas fisik Prodi IKOR FPOK UPI. Adapun Subjek penelitian ini adalah remaja usia 17-21 tahun ( $M \pm SD = 19,34 \pm 0,39$ ) di Jawa Barat, yang terdiri dari 352 remaja laki-laki dan 427 remaja perempuan. Jumlah data yaitu sebanyak

1245, namun setelah dilakukan *screening* diperoleh jumlah data yang dapat diolah dan dianalisis sebanyak 779 data set.

### Steps data

Instrumen yang digunakan yaitu *International Physical Activity Questionnaire* (IPAQ) yang menghasilkan *crawling data* aktivitas selama 7 hari kebelakang. Data diperoleh melalui via internet dengan mengisi google form dan dibagikan melalui aplikasi Whatsapp. Jawaban yang diberikan pada kuesioner berupa waktu aktivitas untuk setiap level yang selanjutnya digunakan untuk menentukan jumlah MET yang sesuai dengan jenis aktivitas dan tingkat atau level intensitas. Skor MET dihasilkan dari kalkulasi waktu aktivitas dengan menggunakan rumus untuk selanjutnya menghasilkan tiga kemungkinan *physical activity level* (PAL) yaitu ‘Berat’, ‘Sedang’, atau ‘Ringan’.

### Mets Data

Metabolic equivalent of task (MET) adalah sebuah unit yang digunakan untuk mengukur intensitas aktivitas fisik seseorang. MET mengacu pada jumlah energi yang dibutuhkan oleh tubuh saat melakukan suatu aktivitas tertentu dibandingkan dengan kebutuhan setiap menit saat beristirahat dengan tenang. Intensitas level MET aktivitas fisik, aktivitas ringan (<3 MET), aktivitas sedang (3-6 MET), aktivitas berat (>6 MET) (Don Hall, 2008).

International Physical Activity Questionnaire (IPAQ), kuesioner yang dirancang terutama untuk pengawasan jenis aktivitas fisik yang dilakukan seseorang dalam kehidupan sehari-hari. Pertanyaan akan menanyakan tentang waktu yang dihabiskan untuk aktif secara fisik di 7 hari terakhir (IPAQ Research Committee, 2004). Rumus yang digunakan IPAQ dalam perhitungan skor aktivitas fisik adalah total aktivitas fisik.

$$\text{Walking MET-minutes/week} = 3.3 * \text{walking minutes} * \text{walking days.}$$
$$\text{Moderate MET-minutes/week} = 4.0 * \text{moderate-intensity activity minutes} * \text{moderate days.}$$
$$\text{Vigorous MET-minutes/week} = 8.0 * \text{vigorous-intensity activity minutes} * \text{vigorous-intensity days}$$
$$\text{Total physical activity MET-minutes/week} = \text{sum of Walking} + \text{Moderate} + \text{Vigorous METminutes/ week scores. (IPAQ Research Committee, 2005).}$$

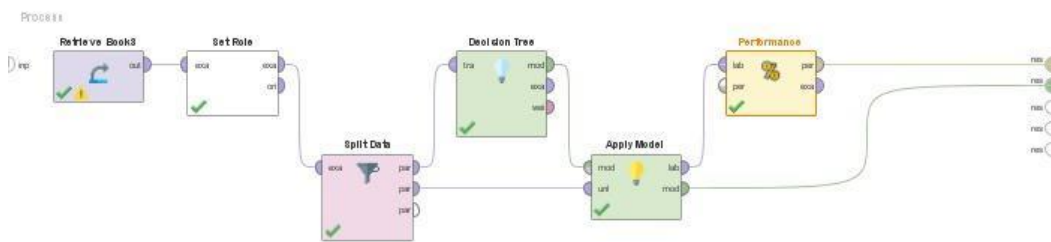
**Tabel 1. Norma IPAQ dan GPAQ**

Kategori	METs (Menit/Minggu)
Berat	$\geq 3000$
Sedang	$\geq 600$
Ringan	$< 600$

---

## Preprocessing

Pada tahap awal yaitu screening data, pemfilteran awal pada tahap ini, jawaban responden yang tidak lengkap, tidak relevan, atau tidak sah dapat diidentifikasi dan dihapus. Kemudian melakukan validasi dan verifikasi, proses ini yaitu pemfilteran jawaban dari responden yang dapat di verifikasi dengan pertanyaan untuk memastikan keakuratan jawaban dan keabsahan data. Tahapan kedua yaitu kodifikasi, proses ini dengan modifikasi jawaban responden dari mengambil informasi data yang diberikan oleh kuesioner, selanjutnya tentukan jumlah METs yang sesuai dengan jenis aktivitas dan tingkat intensitas yang diberikan oleh responden. Kemudian peneliti menguji dengan *machine learning*, model algoritma *machine learning* yang digunakan yaitu *decision tree C45*. Setelah itu melatih model dan menetapkan algoritma pada *machine learning*. Setelah melatih model, evaluasi performa model menggunakan set pengujian. Lalu yang terakhir peneliti dapat mengkategorikan aktivitas fisik tersebut ke dalam tingkat intensitas yang sesuai, seperti rendah, sedang, atau tinggi (Gambar 1).



**Gambar 1. Diagram Proses Pengujian Algoritma *Decision Tree* dengan Rapidminer**

## Analytical Methods

Pada bagian ini peneliti menggunakan metode statistika *decision tree*. *Decision tree* merupakan salah satu metode klasifikasi yang paling populer. Pada prosesnya yaitu pengklasifikasian dari data dan menggambarkan relasi dari variabel atribut  $x$  dan variabel target yang menghasilkan bentuk pohon keputusan diubah menjadi aturan, dan menyederhanakan aturan tersebut (Sari & Mahmudy, 2019) *Decision Tree* adalah sebuah struktur pohon yang mana setiap internal node menyatakan pengujian terhadap suatu atribut, setiap cabang menyatakan output dari pengujian tersebut dan leaf node menyatakan kelas-kelas (Laila, 2021) Salah satu kelebihan *decision tree* yaitu daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat diubah menjadi lebih simpel dan spesifik (Setiawati et al., 2016).

Untuk analisis machine learning digunakan aplikasi RapidMiner (*RapidMiner Studio Educational 10.1.002*). Rapidminer adalah software untuk pengolahan data mining. Yang dilakukan oleh rapidminer text mining yang berkisar dengan analisis teks, mengekstrak pola-pola dari dataset yang besar juga mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan database (Faid et al., 2019).

---

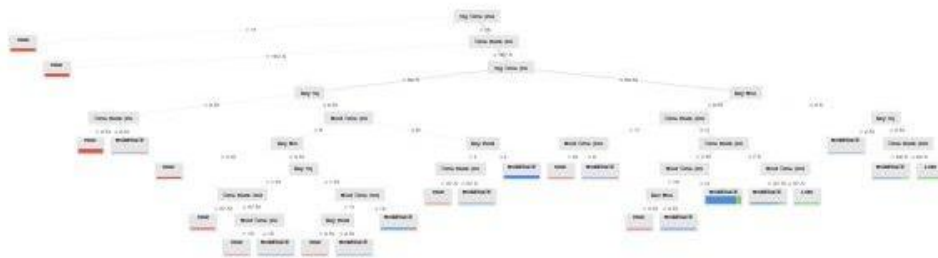
## Train dan test data

Peneliti melatih model menggunakan data yang telah diproses sebelumnya. Jumlah data yaitu sebanyak 1245, namun setelah dilakukan *screening* diperoleh jumlah data yang dapat diolah dan dianalisis sebanyak 779 data set. Adapun jumlah data set untuk training yaitu sebesar 80% dan data untuk test 20% dari keseluruhan data set.

## C. HASIL DAN PEMBAHASAN PENELITIAN

### 1. Hasil

Hasil analisis dengan menggunakan *decision tree* diperoleh pohon keputusan sebagaimana pada gambar 2. Pohon keputusan tersebut menggambarkan hasil data latih analisis klasifikasi, di mana waktu aktivitas moderat menjadi salah satu prediktor utama yang memiliki peran paling besar dalam mengklasifikasi aktivitas fisik.



**Gambar 2. Desain Pohon Keputusan Berdasarkan Waktu Aktivitas Fisik**

Pohon keputusan pada gambar 2 menjelaskan keterkaitan level aktivitas fisik sebagai label dengan variabel atau atribut waktu aktivitas rendah, waktu aktivitas sedang dan waktu aktivitas tinggi. Artinya ketiga atribut tersebut dapat digunakan untuk mengklasifikasi aktivitas fisik. Analisis selanjutnya setelah diperoleh pohon keputusan yaitu menguji model (*apply model*) dan analisis tingkat performa model. Adapun hasilnya sebagaimana pada tabel 2.

Tabel 2. *Confusion Matrix* Berdasarkan Waktu Aktivitas Fisik

	True M	True L	True V	Class Precision
Pred M	<b>65</b>	4	10	82,28%
Pred L	0	<b>9</b>	0	100%
Pred V	6	0	<b>56</b>	90,32%
Recall %	91,5%	69,23%	84,85	

Keterangan:

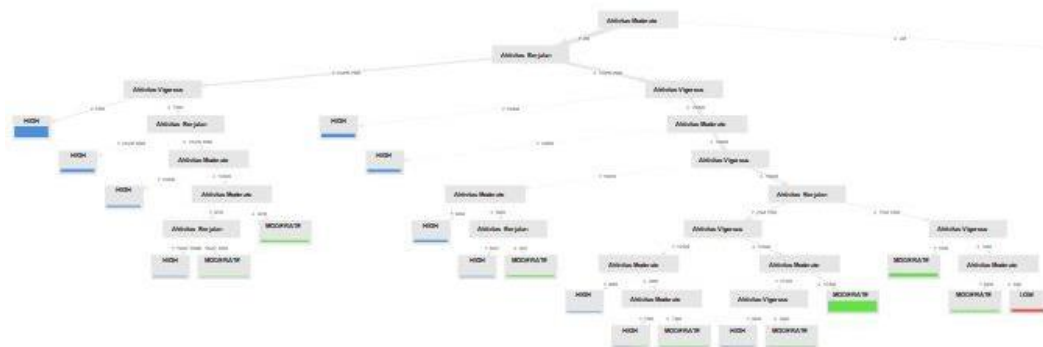
L: Low

M: Moderat

V: Vigorous

Berdasarkan hasil analisis percobaan yang dilakukan terhadap atribut waktu aktivitas fisik sebagaimana pada tabel 2 bahwa (pred M – true M) jumlah data record yang diprediksi termasuk pada klasifikasi aktivitas fisik moderat yaitu sebanyak 65 data sesuai atau benar, sementara 6 data diprediksi termasuk pada klasifikasi aktivitas fisik vigorous. Artinya bahwa algoritma *decision tree* yang digunakan untuk mengklasifikasi aktivitas fisik moderat berdasarkan waktu aktivitas fisik memiliki akurasi sebesar 91,5%. Selanjutnya untuk prediksi klasifikasi aktivitas fisik low (pred L – true L) yaitu sebanyak 9 data benar dan 4 data diprediksi termasuk pada klasifikasi moderat. Tingkat kebenaran untuk kategori aktivitas fisik low yaitu sebesar 69,23 %. Sementara itu untuk pred V – true V sebanyak 56 data benar diklasifikasikan pada aktivitas fisik vigorous dan 10 data diklasifikasikan pada aktivitas fisik moderate dengan kebenaran klasifikasi sebesar 84,85%. Selanjutnya adalah menghitung performa tingkat akurasi dengan analisis *performance vector*. Hasil analisis percobaan jumlah *true positive* (TP) dari M + L + V yaitu  $65 + 9 + 56 = 130$ , artinya bahwa jumlah data *record* yang berhasil diprediksi dengan benar yaitu sebanyak 130 data *record*. Adapun jumlah seluruh data *record* yaitu sebanyak 150 *record*, sehingga dapat dihitung akurasinya yaitu  $(130/150)*100\% = 86,67\%$ .

Analisis kedua dilakukan untuk mengklasifikasikan aktivitas fisik dengan atribut kalkulasi MET pada setiap level. Adapun pohon keputusan dari atribut kalkulasi MET sebagaimana tertera pada gambar 3.



Gambar 3. Desain Pohon Keputusan Berdasarkan kalkulasi MET level

Sama halnya dengan variabel atribut waktu, untuk pemodelan dengan menggunakan dataset MET juga pohon keputusannya (gambar 4) dimulai dari variabel atribut kalkulasi MET aktivitas moderat. Artinya bahwa MET aktivitas moderat menjadi penentu utama dalam mengklasifikasi level aktivitas fisik.

Hasil analisis selanjutnya yaitu berdasarkan data set atribut kalkulasi MET pada setiap klasifikasi aktivitas fisik sebagaimana pada tabel 3 yang merupakan hasil analisis tingkat akurasi performa model *decision tree* berdasarkan dataset MET.

Tabel 3. *Confusion Matrix* Berdasarkan kalkulasi MET level

	True V	True M	True L	Class Precision
Pred V	<b>33</b>	0	0	100%
Pred M	4	<b>32</b>	1	86,49%
Pred L	0	0	<b>7</b>	100%
Recall %	89,19%	100%	87,50%	

Keterangan:

L: Low

M: Moderat

V: Vigorous

Sebagaimana pada tabel 3 bahwa untuk pred V – true V jumlah data record yang diprediksi termasuk pada klasifikasi aktivitas fisik vigorous yaitu sebanyak 33 data sesuai atau benar, sementara 4 data diprediksi termasuk pada klasifikasi aktivitas fisik moderat. Artinya bahwa algoritma *decision tree* yang digunakan untuk mengklasifikasi aktivitas fisik vigorous berdasarkan kalkulasi MET aktivitas fisik memiliki akurasi sebesar 89,19%. Selanjutnya untuk prediksi klasifikasi aktivitas fisik low (pred M – true M) yaitu sebanyak 32 data benar tidak ada data diprediksi termasuk pada klasifikasi vigorous maupun low. Tingkat kebenaran untuk prediksi kategori level aktivitas fisik moderat yaitu sebesar 100 %. Sementara itu untuk pred L – true L sebanyak 7 data record benar diklasifikasikan pada aktivitas fisik low dan 1 data record diklasifikasikan pada aktivitas fisik moderate. Kebenaran prediksi untuk klasifikasi level aktivitas fisik low berdasarkan kalkulasi MET yaitu sebesar 87,50%.

Berdasarkan hasil penghitungan performa tingkat akurasi dengan analisis *performance vector* untuk data record berdasarkan kalkulasi MET diperoleh hasil percobaan jumlah *true positive* (TP) dari V + M + L yaitu  $33 + 32 + 7 = 72$ . Artinya bahwa jumlah data *record* yang berhasil diprediksi dengan benar yaitu sebanyak 72 data *record*. Jumlah seluruh data record yaitu sebanyak 77 record, sehingga dapat dihitung akurasinya yaitu  $(72/77)*100\% = 93,51\%$ .

## Pembahasan

Peneliti ini bertujuan untuk mengklasifikasi aktivitas fisik pada remaja dengan menentukan tingkatan kategori tertentu berbasis sumber data kuesioner dengan menggunakan metode *machine learning* (ML). Hasil penelitian menunjukkan bahwa model machine learning dengan algoritma *decision tree* memiliki akurasi yang baik dalam mengklasifikasi aktivitas fisik. Hasil ini sejalan dengan beberapa penelitian sebelumnya dimana metode machine learning mampu mengklasifikasi berbagai variabel, diantaranya mengklasifikasi aktivitas fisik (Aziz et al., 2021b; Goodlich et al., 2020; Pratama, 2020; Wungo et al., 2023), resiko kejadian (Sihombing & Yulianti,



---

2021), performa akademik, sosial media (Savitri et al., 2021) dan film (Chandani & Wahono, 2015).

Penelitian lainnya yang mengkaji penggunaan algoritma *machine learning* model *decision tree*, *random forest* (RF) dan *support vector machine* (SVM) untuk memprediksi aktivitas fisik menunjukkan bahwa algoritma *support vector machine* memperoleh nilai akurasi tertinggi 90.4%, dan algoritma *decision tree* 84% yang merupakan nilai akurasi terendah (Goodlich et al., 2020). Berbeda halnya dengan penelitian lainnya tentang penggunaan algoritma *machine learning* untuk mengklasifikasi indeks aktivitas fisik dengan sumber data *international physical activity questionnaire*, di mana dari tiga model *machine learning* menunjukkan bahwa model *machine learning decision tree* memiliki tingkat akurasi tertinggi yang mencapai 99% sedangkan *support vector machine* menghasilkan nilai akurasi 89% dengan akurasi paling rendah (Rivera et al., 2022). Adapun penelitian mengenai penggunaan algoritma *machine learning* untuk memprediksi menggunakan metode *global physical activity questionnaire*, dari beberapa model *machine learning* yang digunakan menghasilkan akurasi pengklasifikasian, yang menunjukkan bahwa *random forest* mencapai akurasi 0,921 yang merupakan pengklasifikasi terbaik, diikuti dengan *naïve-Bayes* mencapai akurasi 0,723 yang memiliki nilai akurasi terendah (Rahman et al., 2022).

Dilihat dari pengujian performa menunjukkan bahwa tingkat akurasi ML menggunakan model *decision tree* berbasis dataset atribut MET (93,51%) lebih tinggi dibandingkan dengan dataset waktu aktivitas fisik dengan persentase akurasi 86,67 %. Ini sejalan dengan penelitian sebelumnya yang mengkaji klasifikasi aktivitas fisik menggunakan akselerometer, sensor Girooskop, Kamera, dan GPS menggunakan algoritma *machine learning Ensemble support vector machine* (SVM) dan *support vector machine* (SVM) dengan total data sebanyak 13725 record. Hasilnya menunjukkan *Ensemble support vector machine* (SVM) 87,50% tingkat kepekaan dan *support vector machine* (SVM) 78,79% tingkat kepekaan, dengan performa terbaik dihasilkan ketika membandingkan dataset dengan 80% data latih dan 20% data uji dari total 13725 record karena berhasil meningkatkan akurasi, presisi, dan kepekaan (Wungo et al., 2023). Artinya bahwa jenis data atau variable atribut yang digunakan akan menghasilkan performa akurasi yang berbeda-beda.

Penelitian ini masih terbatas hanya menggunakan satu model pengujian *machine learning* yaitu *decision tree* saja. Pengujian berikutnya diperlukan dengan model algoritma lainnya seperti SVM dan *random forest* dengan atribut yang sama. Sebagaimana beberapa penelitian terdahulu yang menjelaskan bahwa SVM dan *random forest* memiliki akurasi tinggi dalam mengklasifikasi. Sebagaimana penelitian yang membahas klasifikasi aktivitas fisik dengan algoritma *machine learning Ensemble support vector machine* (SVM) yaitu mengklasifikasikan gerakan berjalan, berlari, dan naik tangga berdasarkan sensor accelerometer dan gyroscope pada smartphone bahwa akurasi Support Vector Machine kernel linear sebesar 79.66% dan mengalami

---

peningkatan sebesar 88.01% setelah menggunakan ensemble. Sedangkan akurasi untuk Support Vector Machine kernel RBF sebesar 79.51 dan mengalami peningkatan sebesar 88.04% setelah menggunakan ensemble (Aziz et al., 2021).

#### **D. KESIMPULAN**

Penelitian ini bertujuan menganalisis performa *machine learning decision tree* dalam mengklasifikasi aktivitas fisik berdasarkan waktu dan MET. Hasil penelitian menunjukkan bahwa algoritma decision tree memiliki kemampuan dalam mengklasifikasi aktivitas fisik, di mana pengujian berdasarkan dataset MET memiliki akurasi lebih tinggi dari dataset waktu aktivitas.

#### **DAFTAR PUSTAKA**

- Alexander, L. M., Salum, G. A., Swanson, J. M., & Milham, M. P. (2020). Measuring strengths and weaknesses in dimensional psychiatry. *Journal of Child Psychology and Psychiatry*, *61*(1), 40–50.
- Aziz, F., Usman, S., & Jeffry, J. (2021a). Klasifikasi Physical Activity Berbasis Sensor Accelerometer, Gyroscope, dan Gravity menggunakan Algoritma Multi-class Ensemble GradientBoost. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *5*(4), 1265. <https://doi.org/10.30865/mib.v5i4.3222>
- Besson, H., Brage, S., Jakes, R. W., Ekelund, U., & Wareham, N. J. (2010). Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. *American Journal of Clinical Nutrition*, *91*(1), 106–114. <https://doi.org/10.3945/ajcn.2009.28432>
- Chastin, S. F. M., Culhane, B., & Dall, P. M. (2014). Comparison of self-reported measure of sitting time (IPAQ) with objective measurement (activPAL). *Physiological Measurement*, *35*(11), 2319–2328. <https://doi.org/10.1088/0967-3334/35/11/2319>
- Chandani, V., & Wahono, R. S. (2015). Komparasi Algoritma Klasifikasi Machine Learning Dan Feature Selection pada Analisis Sentimen Review Film. *Journal of Intelligent Systems*, *1*(1). <http://journal.ilmukomputer.org>
- Don Hall, D. C. (2008). *This handout lists the intensity-the MET level-of various physical activities. What level is right for you?*
- Dumith, S. C., Gigante, D. P., Domingues, M. R., & Kohl, H. W. (2011). Physical activity change during adolescence: A systematic review and a pooled analysis. *International Journal of Epidemiology*, *40*(3), 685–698. <https://doi.org/10.1093/ije/dyq272>
- Fahmi Hasan, M., Bahri, S., Sri Ramania, N., Abdul Karim, D., & Dwi Juniarsyah, A. (2019). *TINGKAT AKTIVITAS FISIK SISWA SEKOLAH MENENGAH PERTAMA: Vol. IV* (Issue 2).
- Faid, M., Jasri, M., & Rahmawati, T. (2019). Perbandingan Kinerja Tool Data Mining Weka dan Rapidminer Dalam Algoritma Klasifikasi. *Teknika*, *8*(1), 11–16. <https://doi.org/10.34148/teknika.v8i1.95>

- Goodlich, B. I., Armstrong, E. L., Horan, S. A., Baque, E., Carty, C. P., Ahmadi, M. N., & Trost, S. G. (2020). Machine learning to quantify habitual physical activity in children with cerebral palsy. *Developmental Medicine and Child Neurology*, 62(9), 1054–1060. <https://doi.org/10.1111/dmcn.14560>
- Gustra Jeki, A., Fitria Isnaini, I., Studi Ilmu Gizi, P., Tinggi Ilmu Kesehatan Baiturrahim Jambi, S., & Moh, J. D. (2022). AKTIVITAS FISIK PADA REMAJA DENGAN KEGEMUKAN; SISTEMATIK REVIEW PHYSICAL ACTIVITY IN ADOLESCENT WITH OBESITY; A SYSTEMATIC REVIEW. *Jurnal Ilmu Kesehatan Masyarakat*, 18(2). <https://doi.org/10.19184/ikesma.v18i1.24902>
- Hidayah, F., Pusari, R. W., & Rakhmawati, E. (2019). ANALISA PENGGUNAAN GADGET TERHADAP KESEHATAN MENTAL ANAK USIA DINI. *Seminar Nasional PAUD 2019*, 119–126.
- Hüllermeier, E. (2005). Fuzzy methods in machine learning and data mining: Status and prospects. *Fuzzy Sets and Systems*, 156(3), 387–406.
- IPAQ Research Committee. (2004). *Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ)-Short Form*. [www.ipaq.ki.se](http://www.ipaq.ki.se).
- IPAQ Research Committee. (2005). *Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ)-Short and Long Forms*. [www.ipaq.ki.se](http://www.ipaq.ki.se).
- Marbun, R., Dea, V., Studi Sarjana Terapan Manajemen Informasi Kesehatan, P., Panti Waluya Malang, Stik., & Timur, J. (2021). *EDUKASI KESEHATAN PADA REMAJA DALAM PENTINGNYA GIZI SEIMBANG DAN AKTIVITAS FISIK DI ERA PANDEMI COVID-19*. 4(3).
- Miric, M., Jia, N., & Huang, K. G. (2023). Using supervised machine learning for large-scale classification in management research: The case for identifying artificial intelligence patents. *Strategic Management Journal*, 44(2), 491–519.
- Navada, A., Ansari, A. N., Patil, S., & Sonkamble, B. A. (2011). Overview of use of decision tree algorithms in machine learning. *2011 IEEE Control and System Graduate Research Colloquium*, 37–42.
- Pratama, R. R. (2020). Analisis Model Machine Learning Terhadap Pengenalan Aktifitas Manusia. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 19(2), 302–311. <https://doi.org/10.30812/matrik.v19i2.688>
- Qadrini Laila. (2021). *DECISION TREE DAN ADABOOST PADA KLASIFIKASI PENERIMA PROGRAM BANTUAN SOSIAL*.
- Rahman, H. A., Kwicklis, M., Ottom, M., Amornsriwatanakul, A., Abdul-Mumin, K. H., Rosenberg, M., & Dinov, I. D. (2022). Prediction Modeling of Mental Well-Being Using Health Behavior Data of College Students. *Research Square*. <https://doi.org/10.21203/rs.3.rs-1281305/v1>
- Rivera, O., Castillo-Castaneda, E., Avilés, O. F., & Hernández, R. (2022). Index of Physical activity and Fall Efficacy scale classification through biomechanical signals and Machine Learning. *Journal of Engineering Research*. <https://doi.org/10.36909/jer.16527>
- Rohmah, L., Nur Muhammad S-, H., Jasmani, P., dan Rekreasi, K., & Ilmu Olahraga, F. (2021). *Tingkat Kebugaran Jasmani dan Aktivitas Fisik Siswa Sekolah*. <https://ejournal.unesa.ac.id/index.php/jurnal-pendidikan-jasmani>

- Rondonuwu, R., Tuegeh, J., Bahuwa, S., Sarimin, D. S., Kementrian, P., & Manado, K. (2020). *Aktivitas Fisik dan Penyakit Jantung Koroner*.  
[https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjM\\_oW18Y-AAxXnxjgGHfSrCEgQFnoECBgQAQ&url=https%3A%2F%2Fmedia.hypersites.com%2Fclients%2F1235%2Ffilemanager%2FMHC%2FMETS.pdf&usg=AOvVaw2bLbu-FFnJvzu5hTlcZe6w&opi=89978449](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjM_oW18Y-AAxXnxjgGHfSrCEgQFnoECBgQAQ&url=https%3A%2F%2Fmedia.hypersites.com%2Fclients%2F1235%2Ffilemanager%2FMHC%2FMETS.pdf&usg=AOvVaw2bLbu-FFnJvzu5hTlcZe6w&opi=89978449)
- Sauro, J., & Lewis, J. R. (2011). When designing usability questionnaires, does it hurt to be positive? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2215–2224.
- Savitri, N. L. P. C., Rahman, R. A., Venyutzky, R., & Rakhmawati, N. A. (2021). Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning. *Jurnal Teknik Informatika Dan Sistem Informasi*, 7(1). <https://doi.org/10.28932/jutisi.v7i1.3216>
- Sihombing, P. R., & Yulianti, I. F. (2021). Penerapan Metode Machine Learning dalam Klasifikasi Risiko Kejadian Berat Badan Lahir Rendah di Indonesia. *MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 20(2), 417–426.  
<https://doi.org/10.30812/matrik.v20i2.1174>
- Tan, Z., Yan, Z., & Zhu, G. (2019). Stock selection with random forest: An exploitation of excess return in the Chinese stock market. *Heliyon*, 5(8).
- WHO. (2020). *WHO GUIDELINES ON PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR*.
- WHO. (2021). *Global Physical Activity Questionnaire Analysis Guide GPAQ Analysis Guide Global Physical Activity Questionnaire (GPAQ) Analysis Guide*.  
<http://www.who.int/chp/steps/GPAQ/en/index.html>
- WHO. (2022). *The top 10 causes of death*. [https://www-who-int.translate.goog/news-room/fact-sheets/detail/the-top-10-causes-of-death?\\_x\\_tr\\_sl=en&\\_x\\_tr\\_tl=id&\\_x\\_tr\\_hl=id&\\_x\\_tr\\_pto=tc](https://www-who-int.translate.goog/news-room/fact-sheets/detail/the-top-10-causes-of-death?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc)
- wilda welis, muhamad sazeli rifki. (2013). *BUKU PETUNJUK GIZI UNTUK AKTIFITAS FISIK*.  
<http://repository.unp.ac.id/489/>
- Wungo, S. La, Mardewi, M., Aziz, F., Ishak, P., & SHILI, H. (2023). Classification of Multiclass Ensemble SVM for Human Activities based on Sensor Accelerometer and Gyroscope. *ILKOM Jurnal Ilmiah*, 15(1), 107–117. <https://doi.org/10.33096/ilkom.v15i1.1270.107-117>
- Zwingly C, Manampiring, A. E., V Wariki, W. M., F Palandeng, H. M., & L G Langi, F. F. (2019). :358-367 *Hubungan Kualitas Hidup Kesehatan dengan Aktivitas Fisik dan Status Gizi Remaja di Era Pandemi COVID-19*. 13(3). <https://doi.org/10.35790/jbm.13.3.2021.34417>