

Pengembangan Model Prediksi Indeks Massa Tubuh (BMI) Berbasis Web Menggunakan Metode Decision Tree dengan Algoritma Cart

Abraham Danar Jovian, Aliefian Ramadhan, Muhammad Nabil R, Anggraini Puspita Sari*

Universitas Pembangunan Nasional "Veteran" Jawa Timur

*Penulis korespondensi. E-mail: anggraini.puspita.if@gmail.ac.id

ABSTRACT

This study implements and evaluates the Decision Tree algorithm using the CART (Classification and Regression Tree) method for predicting Body Mass Index (BMI) and develops an interactive website. The primary objective is to determine the effectiveness of the Decision Tree in predicting BMI categories based on gender, height, and weight. BMI data is processed and divided into training (80%) and testing (20%) sets. The Decision Tree model with the CART method is trained and tested, resulting in an accuracy of 89.1%. Evaluation is conducted using a classification report and a confusion matrix. The research outcome is a website that allows users to input their gender, weight, and height to receive a predicted BMI category. This website provides a user-friendly interface, making it easier for the general public to calculate their BMI. This study demonstrates that the Decision Tree with the CART method is effective for BMI prediction and can be used as a tool in the healthcare field for identifying BMI categories.

Keywords

Body Mass Index (BMI),
Decision Tree,
CART,
Classification,
Machine Learning,
Interactive Website.

ABSTRAK

Penelitian ini mengimplementasikan dan mengevaluasi algoritma Decision Tree menggunakan metode CART (Classification and Regression Tree) untuk prediksi Indeks Massa Tubuh (BMI) serta mengembangkan website interaktif. Tujuan utama adalah menentukan efektivitas Decision Tree dalam memprediksi kategori BMI berdasarkan jenis kelamin, tinggi, dan berat badan. Data BMI diolah dan dibagi menjadi set pelatihan (80%) dan pengujian (20%). Model Decision Tree dengan metode CART dilatih dan diuji, menghasilkan akurasi sebesar 89.1%. Evaluasi dilakukan dengan laporan klasifikasi dan matriks kebingungan. Hasil penelitian berupa website yang memungkinkan pengguna memasukkan data gender, berat, dan tinggi badan untuk mendapatkan prediksi kategori BMI. Website ini menyediakan antarmuka user-friendly yang mempermudah masyarakat umum menghitung BMI mereka. Penelitian ini menunjukkan bahwa Decision Tree dengan metode CART efektif untuk prediksi BMI dan dapat digunakan sebagai alat dalam bidang kesehatan untuk identifikasi kategori BMI.

PENDAHULUAN

Indeks Massa Tubuh (BMI) merupakan indikator penting dalam menentukan status gizi seseorang berdasarkan berat dan tinggi badan. Sebagai contoh studi kasus, seorang dokter umum yang ingin memantau status gizi pasiennya dengan cepat dan akurat seringkali dihadapkan pada tantangan perhitungan manual BMI dan klasifikasinya sesuai standar WHO. Proses ini memakan waktu, terutama dengan banyaknya pasien yang harus ditangani. Oleh karena itu, diperlukan alat yang dapat mempermudah proses ini, seperti sistem prediksi otomatis yang dapat mengklasifikasikan BMI secara cepat dan akurat berdasarkan data input sederhana [1][2]. Penelitian ini memiliki beberapa kelebihan utama. Pertama, otomatisasi dan efisiensi yang ditawarkan oleh implementasi algoritma Decision Tree memungkinkan proses klasifikasi BMI dilakukan tanpa perlu perhitungan manual. Kedua, algoritma Decision Tree yang digunakan telah menunjukkan tingkat akurasi yang tinggi dalam memprediksi kategori BMI, mencapai 89.1%. Ketiga, pengembangan website interaktif menyediakan antarmuka yang user friendly, sehingga mudah diakses oleh masyarakat umum tanpa memerlukan pengetahuan teknis mendalam. Keempat, model yang telah dilatih dapat disimpan dan digunakan kembali, memudahkan integrasi dengan sistem lain atau penggunaan di masa mendatang [2][3][4]. Penelitian ini bertujuan untuk mengembangkan model prediksi BMI menggunakan algoritma Decision Tree yang dapat mengklasifikasikan status gizi berdasarkan data jenis kelamin, tinggi badan, dan berat badan.

Algoritma Decision Tree yang digunakan dalam penelitian ini mengimplementasikan metode CART (Classification and Regression Trees), yang merupakan salah satu metode yang paling populer dan efektif dalam pembuatan pohon keputusan. Metode CART bekerja dengan cara membagi dataset menjadi subset-subset yang lebih kecil berdasarkan fitur-fitur input, sehingga menghasilkan model yang mudah diinterpretasi dan akurat dalam klasifikasi [3][5]. Selain itu, penelitian ini juga bertujuan untuk mengevaluasi kinerja model tersebut dengan menggunakan metrik akurasi, laporan klasifikasi, dan matriks kebingungan untuk memastikan efektivitas dan akurasi model [6][7]. Lebih lanjut, penelitian ini mengembangkan sebuah website interaktif yang memungkinkan pengguna untuk memasukkan data pribadi mereka (jenis kelamin, berat badan, dan tinggi badan) dan mendapatkan hasil prediksi BMI secara instan. Dengan demikian, alat bantu ini diharapkan dapat digunakan oleh tenaga kesehatan dan masyarakat umum untuk memonitor status gizi dengan cepat dan akurat [4][8].

METODE

Dalam pelaksanaan penelitian, metode yang dilakukan dibagi menjadi lima tahapan utama. Pertama, mengumpulkan data yang dibutuhkan dari berbagai sumber yang relevan dan terpercaya. Kedua, mempersiapkan data dengan melakukan proses pembersihan dan normalisasi untuk memastikan kualitas data yang optimal. Ketiga, mempersiapkan model CART untuk penerapan terhadap data yang ada, termasuk proses pemilihan parameter dan validasi model. Keempat, dilakukan pembagian dataset menjadi data latih dan data uji, serta pelatihan model untuk pemrosesan dan evaluasi kinerja. Terakhir, masuk kepada perancangan dan pengembangan website interaktif yang digunakan untuk mengaplikasikan hasil penelitian, sehingga pengguna dapat dengan mudah memasukkan data pribadi mereka dan mendapatkan hasil prediksi BMI secara instan.



Gambar 1. Metode Penelitian

Pengumpulan Data

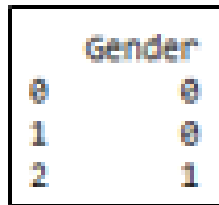
Proses pengumpulan data dimulai dengan memilih dataset yang sesuai dan representatif untuk penelitian ini. Dataset bersumber dari website Kaggle. Kami menggunakan dataset BMI yang mencakup informasi tentang jenis kelamin, tinggi badan, dan berat badan dari berbagai individu. Data ini diakses dari file CSV yang disimpan secara lokal [6][9].

	Gender	Height	Weight	Index
0	0	174	96	4
1	0	189	87	2
2	1	185	110	4
3	1	195	104	3
4	0	149	61	3

Gambar 2. Dataset

Persiapan Data

Agar dataset dapat diproses oleh algoritma pembelajaran mesin, kolom 'Gender' diubah dari teks menjadi nilai numerik: 'Male' diganti dengan 0 dan 'Female' dengan 1. Ini penting untuk mengubah data kategorikal ke format yang bisa digunakan oleh model Klasifikasi Pohon Keputusan [5][10][11].

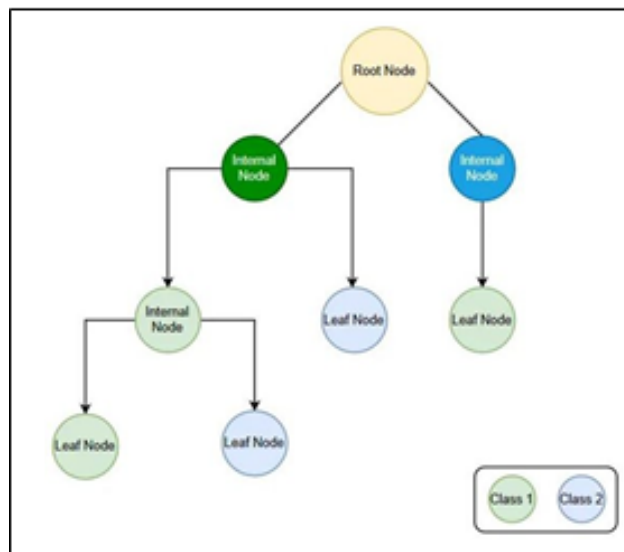


0	0
1	1

Gambar 3. Klasifikasi Gender

Persiapan Model CART

Dalam tahap ini, persiapan model CART (Classification and Regression Tree) melibatkan beberapa langkah penting untuk memastikan model dapat dilatih dan dievaluasi dengan baik. Pertama, dataset BMI yang berisi informasi mengenai gender, tinggi badan, berat badan, dan indeks BMI dibaca dan dimuat ke dalam lingkungan kerja [12][13].



Gambar 4. Model CART

Pembagian dan Pelatihan Data

Dengan menentukan fitur-fitur yang akan digunakan untuk melatih model: 'Gender', 'Height', dan 'Weight'. Sementara itu, variabel targetnya adalah 'Index', yang menunjukkan kategori BMI.

```
x = bmi[['Gender', 'Height', 'Weight']]  
y = bmi['Index']
```

Gambar 5. Indeks Data

Dataset kemudian dibagi menjadi dua bagian: set pelatihan dan set pengujian dengan perbandingan 80:20. Set pengujian (20% dari data) digunakan untuk menguji kinerja model yang telah dilatih. Data ini tidak digunakan selama pelatihan sehingga memberikan indikasi yang lebih

objektif tentang seberapa baik model dapat memprediksi data yang belum pernah dilihat sebelumnya.

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
tree_model = DecisionTreeClassifier()
tree_model.fit(X_train, y_train)
y_pred = tree_model.predict(X_test)
accuracy = round(accuracy_score(y_pred, y_test), 3)
print('Accuracy: ', accuracy)
```

Accuracy: 0.9

Gambar 6. Pelatihan Data

Dengan memisahkan data pengujian dari data pelatihan, kita dapat mengukur sejauh mana model dapat menggeneralisasi pola-pola yang ada dalam data, daripada sekadar menghafal data pelatihan (overfitting). Metode ini dirancang untuk mendukung pembelajaran mesin dalam menghasilkan prediksi yang akurat serta memastikan hasil penelitian yang dapat diandalkan dan valid [8][14][15].

Perancangan Website

Penggunaan operasi prediksi dijalankan melalui website berbasis HTML dan PHP. Dengan merancang operasi PHP untuk mengeksekusi operasi metode prediksi BMI. Lalu, form masukan user di buat berdasar gender, tinggi, dan berat badan. Selanjutnya, membuat output dari operasi yang menampilkan hasil indeks di samping form. Perancangan ini menggunakan program Visual Studio Code dan bahasa pemrograman website (HTML, PHP, CSS) [9].

HASIL DAN PEMBAHASAN

Evaluasi Model

Classification Report:				
	precision	recall	f1-score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	4
2	0.90	1.00	0.95	18
3	0.78	0.88	0.82	8
4	0.96	0.77	0.85	30
5	0.90	0.97	0.94	39
accuracy			0.91	100
macro avg	0.92	0.94	0.93	100
weighted avg	0.91	0.91	0.91	100
Confusion Matrix:				
[[1 0 0 0 0 0]				
[0 4 0 0 0 0]				
[0 0 18 0 0 0]				
[0 0 1 7 0 0]				
[0 0 1 2 23 4]				
[0 0 0 0 1 38]]				

Gambar 7. Model Evaluasi Metriks

Gambar 7 menjelaskan hasil dari evaluasi metriks. Setelah melalui semua proses menghasilkan akurasi dan laporan klasifikasi yang dimana memberikan metrik evaluasi untuk setiap kelas dalam model klasifikasi. Berikut adalah metrik yang disajikan:

- Precision: Proporsi prediksi positif yang benar-benar positif.
- Recall: Proporsi data positif yang benar terprediksi sebagai positif.
- F1-score: Harmonic mean dari precision dan recall.
- Support: Jumlah kejadian sebenarnya dari setiap kelas dalam dataset.

Website



Gambar 8. Tampilan Website

Gambar 8 menampilkan halaman website yang digunakan untuk prediksi oleh pengguna. Pengguna dapat memasukkan gender (laki-laki/perempuan), tinggi (cm), berat (kg). Hasil prediksi BMI akan ditampilkan di samping form. Website ini terdapat beberapa bagian. Bagian pertama berada di sebelah kanan dengan judul "Hasil Prediksi". Terdapat gambar siluet seseorang yang menunjukkan hasil dari prediksi BMI. Pada bagian bawah gambar, terdapat teks yang menunjukkan kategori BMI dari hasil prediksi, yaitu "Kategori BMI: Overweight". Selain itu terdapat saran kesehatan sesuai hasil dari prediksi website.

KESIMPULAN

Penelitian ini berhasil mengembangkan model prediksi Indeks Massa Tubuh (BMI) menggunakan algoritma Decision Tree berbasis metode CART, yang menunjukkan akurasi tinggi hingga 89.1%. Proses penelitian meliputi pengumpulan dan persiapan data dari dataset BMI, konversi data kategorikal, serta pembagian dataset menjadi data latih dan data uji untuk memastikan validitas dan keandalan model. Evaluasi kinerja model dilakukan menggunakan metrik akurasi, laporan klasifikasi, dan matriks kebingungan. Selain itu, penelitian ini juga mencakup pengembangan website interaktif yang memudahkan pengguna dalam memprediksi BMI mereka berdasarkan data jenis kelamin, tinggi badan, dan berat badan. Website ini, dirancang dengan HTML, PHP, dan CSS, menyediakan antarmuka yang user-friendly, memungkinkan pengguna mendapatkan hasil prediksi BMI dan saran kesehatan secara instan, sehingga memudahkan tenaga kesehatan dan masyarakat umum dalam memantau status gizi dengan cepat dan akurat.

DAFTAR PUSTAKA

- [1] Nurussakinah, M., & Faisal, M. (2023). Klasifikasi Penyakit Diabetes Menggunakan Algoritma Decision Tree. *Jurnal Informatika*, 10(2), 143-149
- [2] Juli, A. (2024). Penerapan Algoritma Machine Learning pada Visualisasi Data.
- [3] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees*. CRC Press.
- [4] Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.
- [5] Geron, A. (2019). *Hands-On Machine Learning with ScikitLearn, Keras, and TensorFlow*. O'Reilly Media.
- [6] Boulesteix, A. L., & Strimmer, K. (2007). Partial least squares: a versatile tool for the analysis of highdimensional genomic data. *Briefings in Bioinformatics*, 8(1), 32-44.

- [7] Quinlan, J. R. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann.
- [8] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- [9] Niederst Robbins, J. (2018). *Learning Web Design: A Beginner's Guide to HTML, CSS, JavaScript, and Web Graphics*. O'Reilly Media.
- [10] Sari, A. P., Prasetya, D. A., Al Haromainy, M. M., Aditiawan, F. P., Sihananto, A. N., & Saputra, W. S. (2022, November). Analisis Faktor Kesuksesan Penggunaan eBelajar Menggunakan Metode Hot-Fit di STIKI Malang. In *PROSIDING SEMINAR NASIONAL SAINS DATA* (Vol. 2, No. 1, pp. 92-102).
- [11] Welling, L., & Thomson, L. (2009). *PHP and MySQL Web Development*. 4th ed. Indianapolis, IN: Addison-Wesley.
- [12] Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques*. 3rd ed. San Francisco, CA: Morgan Kaufmann.
- [13] Kurniati, N. M., & Putri, D. A. (2022). Implementasi Algoritma CART untuk Prediksi Klasifikasi Status Gizi. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 9(1), 45-56.
- [14] Rokach, L., & Maimon, O. (2008). *Data Mining with Decision Trees: Theory and Applications*. World Scientific Publishing Company.
- [15] Russell, S., & Norvig, P. (2016). *Artificial Intelligence: A Modern Approach*. 3rd ed. Upper Saddle River, NJ: Pearson