



Desain dan Implementasi Algoritma Decision Tree untuk Klasifikasi Kategori Obesitas

Yusrida Jelianti Sihite¹, Andika Veriando Saragih², Jatmiko Althaf Aziz³

¹²³Matematika dan Ilmu Pengetahuan Alam, Ilmu Komputer, Universitas Negeri Medan, Medan, Indonesia
Email: ¹yusrida041106@gmail.com, ²andikaandikarr@gmail.com, ³althafaziz02@gmail.com

Abstrak— Perkembangan teknologi machine learning memungkinkan pengklasifikasian data kesehatan secara cepat dan akurat. Penelitian ini bertujuan menerapkan algoritma Decision Tree untuk klasifikasi kategori obesitas berdasarkan variabel tinggi badan, berat badan, dan Indeks Massa Tubuh (BMI). Dataset terdiri dari 200 sampel individu yang telah dibersihkan dan diberi label kategori obesitas: Underweight, Normal weight, Overweight, dan Obese. Model Decision Tree dioptimasi dengan pembatasan kedalaman (max_depth=5) untuk mengurangi overfitting dan meningkatkan kemampuan generalisasi. Evaluasi menggunakan data testing menghasilkan akurasi 99,5%, sementara rata-rata cross-validation 5-fold mencapai 0,995. Analisis feature importance menunjukkan bahwa BMI merupakan variabel dominan dalam klasifikasi, sedangkan tinggi badan dan berat badan tidak memberikan kontribusi tambahan. Visualisasi pohon keputusan memperlihatkan jalur klasifikasi yang jelas dan mudah diinterpretasikan, mendukung implementasi model sebagai Decision Support System (DSS) untuk identifikasi awal risiko obesitas. Penelitian ini menegaskan efektivitas Decision Tree dalam klasifikasi kategori obesitas dan konsistensinya dengan literatur sebelumnya, sekaligus memberikan dasar untuk pengembangan model yang lebih kompleks dengan menambahkan fitur pola makan, aktivitas fisik, atau algoritma ensemble.

Kata Kunci: Decision Tree; Klasifikasi; Obesitas; BMI; Machine Learning

Abstract— Advances in machine learning technology enable the rapid and accurate classification of health data. This study aims to apply the Decision Tree algorithm to classify obesity categories based on the variables of height, weight, and Body Mass Index (BMI). The dataset consists of 200 individual samples that have been cleaned and labelled with obesity categories: Underweight, Normal weight, Overweight, and Obese. The Decision Tree model was optimised with a depth constraint (max_depth=5) to reduce overfitting and improve generalisation ability. Evaluation using the test data yielded an accuracy of 99.5%, whilst the average 5-fold cross-validation reached 0.995. Feature importance analysis indicated that BMI is the dominant variable in classification, whereas height and weight do not provide additional contribution. Visualisation of the decision tree reveals clear and easily interpretable classification pathways, supporting the implementation of the model as a Decision Support System (DSS) for the early identification of obesity risk. This study confirms the effectiveness of the Decision Tree in classifying obesity categories and its consistency with previous literature, whilst providing a foundation for the development of more complex models by incorporating dietary patterns, physical activity, or ensemble algorithms.

Keywords: Decision tree; Classification; Obesity; BMI; Machine Learning

1. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong pemanfaatan machine learning secara luas dalam proses pengolahan dan analisis data untuk mendukung pengambilan keputusan. Machine learning memungkinkan sistem komputer mempelajari pola dari data sehingga mampu menghasilkan prediksi atau klasifikasi secara otomatis (Charbuty & Abdulazeez, 2021). Salah satu permasalahan yang sering ditemui dalam machine learning adalah klasifikasi, yaitu proses pengelompokan data ke dalam kategori tertentu berdasarkan pola yang diperoleh dari data pelatihan. Untuk menghasilkan klasifikasi yang tepat, diperlukan algoritma yang mampu mengolah data secara efisien dan memberikan hasil prediksi yang akurat (Yang, 2022).

Machine learning sendiri merupakan cabang dari kecerdasan buatan yang memungkinkan komputer belajar dari data tanpa harus diprogram secara eksplisit. Melalui proses pembelajaran tersebut, sistem dapat mengenali pola dalam data dan menggunakannya untuk membuat prediksi terhadap data baru. Secara umum, metode pembelajaran dalam machine learning meliputi supervised learning, unsupervised learning, dan reinforcement learning. Dalam permasalahan klasifikasi, pendekatan yang paling sering digunakan adalah supervised learning, di mana model



dilatih menggunakan dataset yang telah memiliki label sehingga mampu mengidentifikasi pola yang digunakan untuk mengklasifikasikan data baru (Dirik, 2023).

Penggunaan machine learning saat ini telah diterapkan dalam berbagai bidang, termasuk sektor kesehatan. Salah satu penerapannya adalah dalam analisis kondisi kesehatan, misalnya klasifikasi tingkat obesitas berdasarkan Body Mass Index (BMI). Dengan memanfaatkan teknik machine learning, proses analisis data kesehatan dapat dilakukan dengan lebih cepat dan akurat sehingga dapat membantu dalam proses deteksi dini maupun pengambilan keputusan dalam bidang medis (Kitis & Goker, 2023).

Salah satu algoritma yang cukup populer dalam proses klasifikasi adalah Decision Tree (DT). Algoritma ini merupakan metode pembelajaran berbasis kasus yang membentuk aturan klasifikasi dalam bentuk struktur pohon keputusan berdasarkan kumpulan data yang tersedia. Decision Tree banyak digunakan dalam berbagai penelitian karena mampu menghasilkan model prediksi yang sederhana serta mudah dipahami oleh pengguna (Charbuty & Abdulazeez, 2021).

Struktur Decision Tree terdiri dari node dan cabang, di mana setiap node mewakili atribut yang diuji dan setiap cabang menunjukkan kemungkinan nilai dari atribut tersebut. Proses pembentukan pohon dimulai dari root node dan berlanjut hingga mencapai leaf node, yang berfungsi sebagai hasil akhir dari proses klasifikasi. Keunggulan utama metode ini terletak pada kemampuannya dalam memberikan aturan keputusan yang jelas serta memvisualisasikan proses pengambilan keputusan secara sistematis (Blockeel et al., 2023).

Dalam proses pembelajarannya, Decision Tree umumnya melibatkan dua tahapan utama. Tahap pertama adalah membangun struktur pohon menggunakan dataset pelatihan sehingga model dapat mempelajari pola hubungan antara atribut dan kelas target. Tahap kedua adalah menggunakan model yang telah terbentuk untuk melakukan klasifikasi terhadap data baru yang belum pernah diproses sebelumnya. Dengan pendekatan ini, model mampu menghasilkan prediksi berdasarkan aturan yang telah dipelajari dari data pelatihan.

Selain memberikan hasil klasifikasi, Decision Tree juga memungkinkan proses pengambilan keputusan divisualisasikan dalam bentuk grafik yang terdiri dari node keputusan, node peluang, serta node hasil akhir. Representasi visual tersebut membuat proses analisis menjadi lebih mudah dipahami dibandingkan beberapa algoritma machine learning lain yang memiliki struktur model yang lebih kompleks.

Berbagai penelitian sebelumnya menunjukkan bahwa algoritma Decision Tree memiliki kinerja yang baik dalam berbagai permasalahan klasifikasi, termasuk pada bidang kesehatan. Beberapa studi melaporkan bahwa metode ini mampu menghasilkan tingkat akurasi yang tinggi serta memberikan pemahaman yang lebih jelas mengenai faktor-faktor yang mempengaruhi hasil klasifikasi.

Penelitian yang dilakukan oleh Sarhan (2023) menunjukkan bahwa algoritma Decision Tree mampu mengklasifikasikan tingkat obesitas dengan tingkat akurasi sebesar 94,16%, dengan nilai precision 93,93%, sensitivity 96,10%, dan specificity 91,66%. Hasil tersebut menunjukkan bahwa model mampu membedakan kategori obesitas dengan baik berdasarkan faktor-faktor yang mempengaruhi kondisi kesehatan individu (Enhyymayed & Mahmood, 2023).

Penelitian lainnya juga membandingkan beberapa algoritma machine learning seperti Random Forest, XGBoost, CatBoost, dan Artificial Neural Network (ANN) dalam proses klasifikasi obesitas. Hasil penelitian menunjukkan bahwa algoritma berbasis pohon keputusan secara umum memiliki performa yang lebih baik dibandingkan metode lainnya, baik dari sisi akurasi maupun efisiensi waktu pelatihan (Citak & Kolklu, 2025).

Selain itu, penelitian yang membandingkan algoritma Decision Tree dengan Naive Bayes menunjukkan bahwa Decision Tree mampu menghasilkan akurasi yang lebih tinggi, yaitu sebesar 97,47%, dibandingkan dengan Naive Bayes yang mencapai 94,94%. Temuan ini menunjukkan bahwa Decision Tree merupakan metode yang efektif untuk menangani permasalahan klasifikasi multi-kelas serta mampu mengolah berbagai jenis atribut data (Garba & Wurnor, 2022).

Berdasarkan berbagai penelitian tersebut, algoritma Decision Tree memiliki keunggulan dalam hal akurasi, kemudahan interpretasi, serta kemampuan visualisasi model klasifikasi. Oleh karena itu, penelitian ini menggunakan algoritma Decision Tree untuk melakukan klasifikasi kategori obesitas serta menganalisis pola keputusan yang dihasilkan oleh model yang dibangun.

2. METODE

Metode penelitian ini mencakup sumber data, teknik pengumpulan data, model penelitian, definisi operasional variabel, serta metode analisis data yang digunakan untuk mengimplementasikan algoritma Decision Tree dalam klasifikasi kategori obesitas.

2.1 Data dan Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini berupa dataset yang berisi informasi karakteristik fisik individu, yaitu tinggi badan (Height), berat badan (Weight), Body Mass Index (BMI), serta kategori obesitas (ObesityCategory). Dataset diperoleh dari sumber data publik yang tersedia secara terbuka dan telah melalui proses pembersihan data (*data cleaning*) sebelum digunakan dalam proses analisis. Dalam penelitian ini, kategori obesitas digunakan sebagai variabel target yang akan diprediksi oleh model. Sementara itu, variabel Height, Weight, dan BMI digunakan sebagai variabel prediktor yang menjadi dasar dalam proses pembentukan model klasifikasi menggunakan algoritma Decision Tree.

Tabel 3.1. Contoh Data Awal

Height	Weight	BMI	ObesityCategory
173.57	71.98	23.89	Normal weight
164.12	89.95	33.39	Obese
168.07	72.93	25.81	Overweight
168.45	84.88	29.91	Overweight
183.56	69.03	20.48	Normal weight

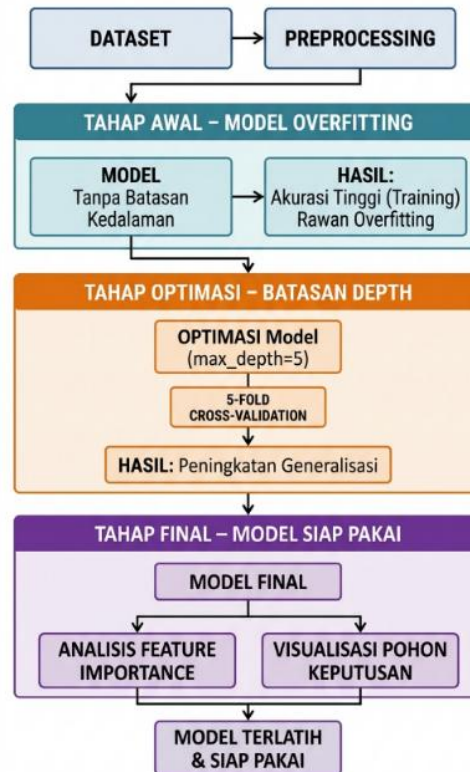
Keterangan: Data di atas adalah lima sampel awal dari dataset lengkap.

2.2. Model Penelitian

Penelitian ini menerapkan algoritma **Decision Tree Classifier** sebagai model utama dalam proses klasifikasi kategori obesitas. Tahapan penelitian dilakukan melalui tiga tahap utama yang bertujuan untuk membangun, mengoptimasi, serta mengevaluasi model klasifikasi yang dihasilkan.

1. **Tahap Awal – Model Overfitting**, Pada tahap pertama, model Decision Tree dibangun tanpa pembatasan kedalaman pohon. Tahap ini bertujuan untuk memahami pola awal yang terdapat pada dataset serta melihat bagaimana model mempelajari karakteristik data, meskipun kondisi tersebut berpotensi menimbulkan overfitting.
2. **Tahap Optimasi – Pembatasan Kedalaman Pohon**, Pada tahap berikutnya dilakukan proses optimasi dengan membatasi kedalaman pohon menggunakan parameter `max_depth = 5`. Pembatasan ini bertujuan untuk meningkatkan kemampuan generalisasi model sehingga tidak terlalu menyesuaikan diri dengan data pelatihan. Evaluasi performa model dilakukan menggunakan metode *5-fold cross-validation*.
3. **Tahap Akhir – Model Klasifikasi**, Pada tahap akhir, model Decision Tree yang telah dioptimasi digunakan untuk melakukan klasifikasi kategori obesitas. Selain itu, dilakukan analisis feature importance untuk mengetahui fitur yang paling berpengaruh terhadap hasil klasifikasi, serta visualisasi pohon keputusan untuk mempermudah interpretasi model.

IMPLEMENTASI DECISION TREE UNTUK KLASIFIKASI KATEGORI OBESITAS



Gambar 2.1. Alur Implementasi Decision Tree

2.3. Algoritma Decision Tree

Algoritma Decision Tree membangun model klasifikasi dalam bentuk struktur pohon yang terdiri dari root node, decision node, dan leaf node. Proses pembentukan pohon dilakukan secara top-down, di mana pada setiap node dipilih atribut terbaik yang mampu memisahkan data secara optimal.

Pemilihan atribut terbaik dilakukan menggunakan konsep Entropy dan Information Gain. Entropy digunakan untuk mengukur tingkat ketidakpastian atau ketidakteraturan dalam suatu dataset, sedangkan Information Gain digunakan untuk menentukan atribut yang memberikan pengurangan entropy terbesar setelah data dipisahkan.

Rumus Entropy dalam algoritma Decision Tree adalah:

$$E(S) = - \sum_{i=1}^c p_i \log_2(p_i)$$

Penjelasan:

$E(S)$ = nilai entropy dari dataset S

p_i = probabilitas kemunculan kelas ke- i dalam dataset

c = jumlah kelas dalam dataset

$\log_2$ = logaritma basis 2



Selanjutnya, nilai Information Gain digunakan untuk menentukan atribut yang paling efektif dalam proses pemisahan data. Perhitungan Information Gain dapat dinyatakan dengan persamaan berikut:

$$Gain(S, A) = E(S) - \sum_{v \in values(A)} \frac{|S_v|}{|S|} E(S_v)$$

di mana:

$Gain(S, A)$ adalah nilai information gain dari atribut A

S adalah dataset awal

S_v adalah subset data dengan nilai atribut tertentu

A adalah atribut yang digunakan untuk pemisahan data

Atribut dengan nilai information gain tertinggi akan dipilih sebagai node berikutnya dalam pembentukan pohon keputusan. Dalam beberapa pendekatan berbasis probabilitas, proses seleksi fitur juga dapat dinyatakan menggunakan persamaan berikut:

$$s = (s_1 + s_2 + s_3 + s_4 + s_5) \bmod (p - 1)$$

di mana $s_1, s_2, \dots, s_5$ merupakan nilai atribut terpilih dan p merupakan nilai modulus yang digunakan untuk normalisasi probabilitas.

2.4. Metode Analisis Data

Setelah model Decision Tree berhasil dibangun, tahap selanjutnya adalah melakukan analisis untuk menilai kinerja model dalam mengklasifikasikan kategori obesitas serta mengetahui pengaruh masing-masing fitur terhadap hasil prediksi. Penilaian performa model dilakukan menggunakan beberapa metrik evaluasi klasifikasi, yaitu accuracy, precision, recall, serta F1-score yang digunakan untuk melihat keseimbangan antara nilai precision dan recall.

Selain itu, penelitian ini juga menerapkan metode 5-fold cross-validation untuk menguji kestabilan model. Teknik ini bertujuan memastikan bahwa hasil evaluasi tidak bergantung pada satu pembagian data tertentu sehingga performa model dapat dinilai secara lebih objektif. Analisis feature importance juga dilakukan untuk mengetahui fitur yang memiliki kontribusi terbesar dalam proses klasifikasi kategori obesitas, sehingga dapat memberikan pemahaman lebih lanjut mengenai faktor yang mempengaruhi hasil prediksi.

Hasil analisis kemudian disajikan melalui visualisasi pohon keputusan (decision tree) serta diagram batang feature importance. Visualisasi tersebut membantu pembaca dalam memahami alur proses klasifikasi dan melihat kontribusi setiap fitur secara lebih jelas. Dengan pendekatan ini, model yang dihasilkan tidak hanya memiliki tingkat akurasi yang baik, tetapi juga mudah diinterpretasikan serta berpotensi digunakan sebagai Decision Support System (DSS) dalam memprediksi kategori obesitas pada dataset nyata.

3. ANALISA DAN PEMBAHASAN

3.1 Hasil Model Decision Tree

Penelitian ini dilakukan dengan tiga tahap model Decision Tree. Tahap awal membangun pohon keputusan tanpa batasan kedalaman (depth), sehingga akurasi pada data training mencapai 100%, tetapi rawan overfitting. Hal ini ditunjukkan dari hasil prediksi data testing yang juga menunjukkan akurasi tinggi pada dataset kecil namun model belum generalisasi untuk data baru.

Tahap optimasi dilakukan dengan membatasi max_depth=5 untuk mengurangi overfitting. Evaluasi dengan data testing menghasilkan akurasi sebesar 99,5%, sedangkan cross-validation 5-fold menunjukkan rata-rata akurasi 0,995. Hasil ini menegaskan bahwa model mampu

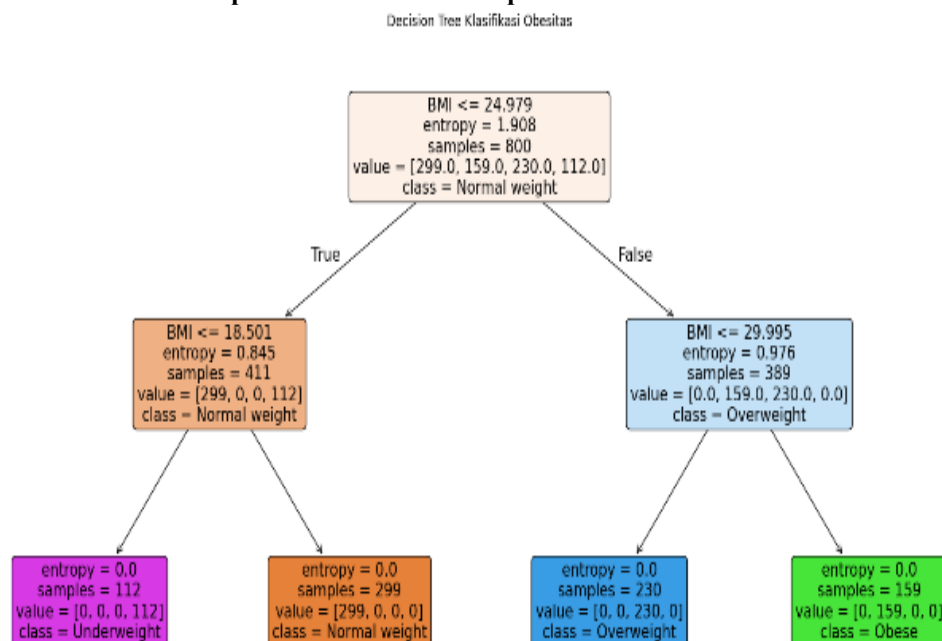
menggeneralisasi dengan baik dan stabil (perhatikan juga font yang digunakan pada Tabel 1). Tabel berikut menunjukkan jumlah sampel per kategori obesitas dan hasil prediksi model:

Tabel 4.1. Distribusi kategori obesitas dan evaluasi model Decision Tree

<i>Kategori Obesitas</i>	<i>Support</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Normal weight	72	1.00	1.00	1.00
Obese	32	1.00	1.00	1.00
Overweight	65	1.00	1.00	1.00
Underweight	31	1.00	1.00	1.00
Average / Total	200	1.00	1.00	1.00

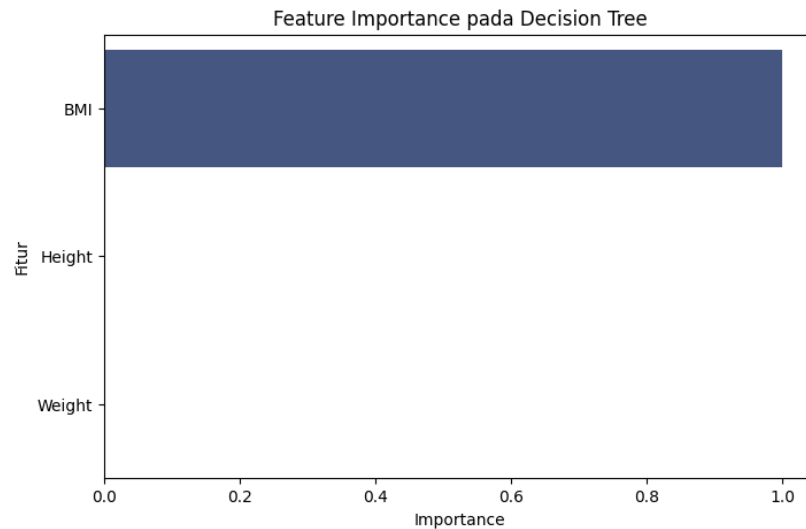
Hasil menunjukkan bahwa model Decision Tree yang dioptimasi mampu mengklasifikasikan setiap kategori obesitas dengan sangat baik, bahkan untuk kelas yang jumlah sampelnya lebih sedikit seperti Underweight.

3.2 Visualisasi Pohon Keputusan dan Feature Importance



Gambar 3.1. Decision Tree Klasifikasi Obesitas

Gambar di atas menampilkan visualisasi Decision Tree yang digunakan untuk mengklasifikasi kategori berat badan berdasarkan Indeks Massa Tubuh (BMI). Pohon keputusan dimulai dari simpul akar (root node) yang mengevaluasi apakah $BMI \leq 24,979$. Jika ya, alur bergerak ke cabang kiri; jika tidak, alur bergerak ke cabang kanan. Pada cabang kiri, model memeriksa apakah $BMI \leq 18,501$ untuk membedakan antara kategori Underweight dan Normal weight, sedangkan pada cabang kanan, model membedakan antara Overweight dan Obese dengan ambang batas $BMI \leq 29,995$. Setiap kotak pada pohon menampilkan informasi tambahan berupa nilai entropy, jumlah sampel, distribusi sampel untuk tiap kategori (Underweight, Normal, Overweight, Obese), serta prediksi akhir dari simpul tersebut. Entropy 0,0 pada daun menunjukkan klasifikasi yang murni dan sempurna.



Gambar 3.2. Feature Importance pada Decision Tree

Selain itu, analisis *feature importance* menunjukkan bahwa variabel BMI memiliki kontribusi penuh sebesar 1,0, menandakan bahwa model sepenuhnya mengandalkan BMI dalam proses klasifikasi. Variabel lain seperti *Height* dan *Weight* memiliki nilai 0,0, karena setelah adanya BMI, kedua fitur ini dianggap tidak memberikan informasi tambahan yang signifikan. Hasil ini menegaskan bahwa pohon keputusan sangat efisien namun bergantung pada satu indikator utama. Dalam konteks medis, BMI terbukti sebagai prediktor paling kuat untuk klasifikasi kategori obesitas. Temuan ini konsisten dengan literatur sebelumnya dan menunjukkan bahwa model ini dapat digunakan sebagai Decision Support System (DSS) untuk prediksi awal kategori obesitas secara cepat dan akurat.

3.3 Implementasi Python dan Analisis Model

Penelitian ini menggunakan bahasa pemrograman Python untuk membangun, melatih, dan memvisualisasikan model Decision Tree dalam klasifikasi kategori obesitas. Langkah awal meliputi pemisahan variabel fitur dan target, pengkodean label kategori, serta pembagian dataset menjadi data training dan testing. Variabel Height, Weight, dan BMI digunakan sebagai fitur numerik, sedangkan ObesityCategory menjadi target variabel. Label encoding dilakukan agar model dapat mengenali kategori kelas sebagai nilai numerik. Dataset kemudian dibagi dengan proporsi 80% untuk training dan 20% untuk testing.

```
x = data[['Height', 'Weight', 'BMI']]
y = data['ObesityCategory']

le = LabelEncoder()
y_encoded = le.fit_transform(y)

x_train, x_test, y_train, y_test = train_test_split(
    x, y_encoded, test_size=0.2, random_state=42
)
```

Gambar 3.3 Kode Persiapan Data

Setelah data siap, model Decision Tree dibangun dalam dua tahap. Tahap pertama adalah model awal tanpa batasan kedalaman, untuk memahami karakteristik dataset meskipun rawan overfitting. Tahap kedua adalah model dioptimasi dengan `max_depth=5`, sehingga mampu menggeneralisasi lebih baik. Evaluasi dilakukan menggunakan accuracy, classification report, dan cross-validation 5-fold, menghasilkan akurasi tinggi pada data testing serta rata-rata CV sebesar 0,995. Di tengah-tengah penjelasan ini, pohon keputusan divisualisasikan untuk memberikan pemahaman intuitif jalur klasifikasi:

```
plt.figure(figsize=(20,10))
plot_tree(
    dt_optimized,
    feature_names=X.columns,
    class_names=le.classes_,
    filled=True,
    rounded=True
)
plt.title("Decision Tree Klasifikasi Obesitas")
plt.show()
```

Gambar 3.4 Kode Visualisasi Decision Tree

Pohon keputusan ini menampilkan jalur klasifikasi dari root node hingga leaf node, lengkap dengan jumlah sampel, distribusi kategori, dan prediksi akhir. Entropy 0,0 pada daun menunjukkan klasifikasi murni, menandakan model mampu membedakan kategori obesitas dengan jelas. Selanjutnya, analisis feature importance dilakukan untuk menilai kontribusi setiap fitur terhadap prediksi. Hasilnya divisualisasikan menggunakan diagram batang di tengah penjelasan:

```
feature_importance = dt_optimized.feature_importances_
fi_df = pd.DataFrame({
    'Fitur': X.columns,
    'Importance': feature_importance
}).sort_values(by='Importance', ascending=False)

plt.figure(figsize=(8,5))
sns.barplot(x='Importance', y='Fitur', data=fi_df, palette="viridis")
plt.title("Feature Importance pada Decision Tree")
plt.show()
```

Gambar 3.5 Kode Diagram Feature Importance

Analisis menunjukkan bahwa BMI memiliki kontribusi penuh sebesar 1,0, sedangkan Height dan Weight tidak memberikan informasi tambahan signifikan. Visualisasi ini mempermudah interpretasi jalur klasifikasi dan mendukung implementasi model sebagai Decision Support System (DSS) untuk prediksi awal kategori obesitas. Dengan demikian, pembaca dapat langsung melihat visualisasi pohon dan kontribusi fitur tepat saat membaca penjelasan, tanpa menunggu sampai akhir.

4. KESIMPULAN

Berdasarkan hasil analisis menggunakan algoritma Decision Tree, model yang dioptimasi dengan parameter *max_depth* = 5 mampu mengklasifikasikan kategori obesitas (Underweight, Normal weight, Overweight, dan Obese) dengan tingkat akurasi yang sangat tinggi, yaitu 99,5% pada data pengujian dan rata-rata 0,995 pada *5-fold cross-validation*. Hasil analisis *feature importance* menunjukkan bahwa BMI merupakan variabel paling dominan dalam menentukan kategori obesitas, sedangkan tinggi badan dan berat badan tidak memberikan kontribusi tambahan yang signifikan setelah BMI digunakan dalam model. Visualisasi pohon keputusan juga memberikan jalur klasifikasi yang jelas sehingga model tidak hanya akurat tetapi juga mudah diinterpretasikan dan berpotensi digunakan sebagai Decision Support System (DSS) untuk prediksi awal kategori obesitas.

Namun demikian, penelitian ini masih memiliki keterbatasan pada jumlah dan jenis variabel yang digunakan, sehingga penelitian selanjutnya disarankan untuk menambahkan fitur lain seperti pola makan, aktivitas fisik, serta membandingkan performa dengan algoritma lain seperti Random Forest atau XGBoost untuk meningkatkan kemampuan prediksi model.



REFERENCES

- Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of applied science and technology trends*, 2(01), 20-28. <https://doi.org/10.38094/jastt20165>
- Yang, B. (2022). Clothing design style recommendation using decision tree algorithm combined with deep learning. *Computational intelligence and neuroscience*, 2022(1), 5745457. <https://doi.org/10.1155/2022/5745457>
- Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of applied science and technology trends*, 2(01), 20-28.
- Guo, Y. (2021). University classroom teaching model based on decision tree analysis and machine learning. *Mobile Information Systems*, 2021(1), 6926013. <https://doi.org/10.38094/jastt20165>
- Dirik, M. (2023). Application of machine learning techniques for obesity prediction: a comparative study. *Journal of complexity in Health Sciences*, 6(2), 16-34.
- Kitis, S., & Goker, H. (2023). Detection of obesity stages using machine learning algorithms. *Anbar Journal of Engineering Sciences*, 14(1), 80-88. <https://doi.org/10.37649/aengs.2023.139350.1045>
- Airlangga, G. (2025). A Comparative Analysis of Machine Learning Models for Obesity Prediction. *Jurnal Informatika Ekonomi Bisnis*, 1-5. <https://doi.org/10.37034/infec.v7i1.1089>
- Blockeel, H., Devos, L., Frénay, B., Nanfack, G., & Nijssen, S. (2023). Decision trees: from efficient prediction to responsible AI. *Frontiers in artificial intelligence*, 6, 1124553. <https://doi.org/10.3389/frai.2023.1124553>
- Mahbooba, B., Timilsina, M., Sahal, R., & Serrano, M. (2021). Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. *Complexity*, 2021(1), 6634811. <https://doi.org/10.1155/2021/6634811>
- Alalawi, H. H., & Alsuwat, M. S. (2021). Detection of cardiovascular disease using machine learning classification models. *International Journal of Engineering Research & Technology*, 10(7), 151-7.
- Ehymmayed, H. M., & Mahmood, S. E. (2023). Decision tree in Obesity Level Classification of Northern Technical University Students. *NTU Journal of Agriculture and Veterinary Science*, 3(4). <https://doi.org/10.56286/ntujavs.v3i4.707>
- CITAK, A. C., BAYRAM, S. S., & KOKLU, M. (2025). Classification of Obesity Levels Using Machine Learning Algorithms. *Intelligent Methods In Engineering Sciences*, 4(3), 100-113. <https://doi.org/10.58190/imiens.2025.157>
- Garba, S., Abdullahi, M., Umar, U. A., & Wurnor, N. T. (2022). Obesity level classification based on decision tree and naïve Bayes classifiers. *SLU Journal of Science and Technology*, 3(1 & 2), 113-121.