

Klasifikasi Obesitas Menggunakan *Domain-Based Decision Tree*, *XGBoost*, dan SHAP

Embay Rohaeti¹ *, Ani Andriyati², Muhammad Edy Rizal³

^{1,2}Program Studi Matematika, Universitas Pakuan, Bogor, Indonesia

³Program Studi Sains Data, Universitas Tadulako, Palu, Indonesia
embay.rohaeti@unpak.ac.id

Diajukan: 16 Agustus 2025, Diperbaiki: 25 Januari 2026, Diterima: 26 Januari 2026

Abstrak

Obesitas merupakan masalah kesehatan masyarakat yang kompleks dan memerlukan strategi identifikasi dini yang efektif untuk mengurangi dampak kesehatan jangka panjangnya. Penelitian ini bertujuan untuk mengklasifikasikan tingkat obesitas, yang dikategorikan sebagai *Underweight*, *Normal*, *Overweight*, dan *Obese*, dengan menggunakan 14 variabel yang dikelompokkan ke dalam tiga domain: biologis, pola makan, dan aktivitas fisik. Selain mengembangkan model prediksi yang akurat, penelitian ini juga mengkaji domain mana yang paling berkontribusi terhadap klasifikasi obesitas. Kami menggunakan dua strategi pemodelan yang saling melengkapi, yaitu model *decision tree* berbasis domain untuk mengevaluasi kekuatan prediktif setiap domain secara independen, dan model komprehensif menggunakan *eXtreme Gradient Boosting* (XGBoost) yang dilatih pada seluruh variabel. Untuk mengatasi ketidakseimbangan kelas, dilakukan *SMOTENC oversampling* pada data latih, dan penyetelan *hyperparameter* dilakukan melalui *cross-validation* pada kedua pendekatan. Evaluasi pada data uji menunjukkan bahwa model XGBoost mengungguli *decision tree* berbasis domain pada semua metrik kinerja, termasuk *balanced accuracy*, *precision*, *recall*, *specificity*, dan *F1-score*. Meskipun *decision tree* menawarkan kemudahan interpretasi pada tingkat domain, *decision tree* tidak memiliki akurasi prediksi seperti model terintegrasi. Analisis SHAP (*SHapley Additive Explanations*) menunjukkan bahwa fitur-fitur yang memengaruhi klasifikasi obesitas tersebar di semua domain, dengan variabel usia, frekuensi konsumsi sayur, dan jenis transportasi sebagai variabel utama. Temuan ini menunjukkan bahwa integrasi data multi-domain meningkatkan akurasi sekaligus kemudahan interpretasi model klasifikasi obesitas, sehingga mendukung penggunaan *interpretable machine learning* untuk penilaian risiko kesehatan personal dan strategi pencegahan.

Kata Kunci: Obesitas; Machine Learning; XGBoost; SHAP; SMOTENC

Abstract

Obesity is a complex public health issue that requires effective early identification strategies to mitigate its long-term health impacts. This study aimed to classify obesity levels, categorized as *Underweight*, *Normal*, *Overweight*, and *Obese*, using 14 predictors grouped into three domains: biological, dietary, and physical activity. In addition to developing an accurate predictive model, the study investigated which domain contributes most to obesity classification. Two complementary modeling strategies were applied: domain-specific decision tree models to evaluate the predictive strength of each domain independently, and a comprehensive model using *eXtreme Gradient Boosting* (XGBoost) trained on all predictors. To address class imbalance, *SMOTENC oversampling* was applied to the training set, and *hyperparameter tuning* was performed via *cross-validation* for both approaches. Evaluation on the test set showed that the XGBoost model outperformed the domain-based decision trees across all performance metrics, including *balanced accuracy*, *precision*, *recall*, *specificity*, and *F1-score*. While decision trees offered domain-level interpretability, they lacked the predictive power of the integrated model. SHAP (*SHapley Additive Explanations*) analysis revealed that influential features spanned all domains, with Age,

Vegetable consumption, and Transportation type emerging as top predictors. These findings demonstrate that integrating multi-domain behavioral data enhances both the accuracy and interpretability of obesity classification models, supporting the use of interpretable machine learning for personalized health risk assessment and prevention strategies.

Keywords: Obesity, Machine Learning, XGBoost, SHAP, SMOTENC

1 Pendahuluan

Obesitas merupakan masalah kesehatan global yang semakin berkembang dan menimbulkan risiko signifikan bagi individu maupun masyarakat [1]. Kondisi ini berkaitan dengan peningkatan risiko penyakit kronis seperti diabetes tipe 2, penyakit kardiovaskular, dan beberapa jenis kanker, serta berkontribusi pada peningkatan biaya perawatan kesehatan dan penurunan kualitas hidup [2]. *World Health Organization* (WHO) melaporkan bahwa tingkat obesitas global telah meningkat secara drastis sejak tahun 1975, dan masalah ini tidak lagi terbatas pada negara berpenghasilan tinggi. Obesitas kini semakin memengaruhi populasi berpenghasilan menengah dan rendah, termasuk remaja dan anak-anak [3], [4], [5]. Mengingat etiologi yang kompleks dan dampak yang luas, strategi identifikasi dini dan pencegahan menjadi penting dalam mengatasi epidemi obesitas.

Machine learning dapat mendukung deteksi dini obesitas dengan menganalisis data yang beragam, mencakup aspek perilaku, biologis, dan gaya hidup. Studi terkini telah menerapkan model seperti *decision tree*, *random forest*, *support vector machine*, dan XGBoost untuk mengklasifikasikan individu ke dalam kategori risiko obesitas berdasarkan variabel kesehatan [6], [7], [8], [9]. Meskipun model-model ini akurat, kemudahan interpretasinya sering kali terabaikan, padahal interpretasi penting bagi pengambilan keputusan dalam kesehatan masyarakat. Selain itu, penelitian yang ada jarang mengkaji kategori variabel yang paling berkontribusi dalam prediksi obesitas sehingga membatasi potensi intervensi yang lebih terarah.

Berbeda dengan penelitian tersebut, studi ini menggabungkan pemodelan berbasis domain dengan analisis interpretabilitas SHAP untuk menilai pengaruh relatif variabel dari sisi biologis, pola makan, dan aktivitas fisik. Kebaruan penelitian ini terletak pada integrasi pendekatan multi-domain dan analisis SHAP untuk menghasilkan model yang tidak hanya akurat tetapi juga mudah diinterpretasi dan relevan untuk pengambilan keputusan.

Kami mengusulkan strategi pemodelan dengan dua tahapan. Pertama, kami melatih model *decision tree* berbasis domain menggunakan variabel yang dikelompokkan ke dalam tiga domain: biologis, pola makan, dan aktivitas fisik. Pendekatan ini memudahkan interpretasi pada masing-masing domain secara terpisah. Kedua, kami melatih model XGBoost yang memanfaatkan seluruh variabel yang tersedia untuk memaksimalkan kinerja prediksi. Untuk meningkatkan kemudahan

interpretasi, analisis SHAP digunakan untuk mengevaluasi tingkat kepentingan fitur baik pada tingkat variabel individu maupun tingkat domain yang lebih luas.

Tujuan utama penelitian ini adalah mengidentifikasi secara jelas jenis data perilaku yang paling efektif dalam memprediksi tingkat obesitas, serta menunjukkan bagaimana *interpretable machine learning model* dapat dimanfaatkan untuk melakukan prediksi secara akurat sekaligus menghasilkan informasi yang dapat ditindaklanjuti. Dengan mengintegrasikan model prediktif dan pemodelan berbasis domain, penelitian ini berkontribusi pada penilaian risiko dan penyusunan strategi pencegahan obesitas yang lebih terarah. Pendekatan ini tidak hanya memprediksi obesitas tetapi juga menjelaskan faktor-faktor yang mendasarinya.

2 Metode Penelitian

Data yang digunakan dalam penelitian ini diperoleh dari [8], berisi 1.610 responden di Türkiye dengan 14 variabel prediktor dan satu variabel target multikelas yang merepresentasikan status obesitas. Variabel prediktor dikelompokkan ke dalam tiga domain: biologis, perilaku makan, dan aktivitas fisik. Variabel target (Kelas) terdiri atas empat kategori: *Underweight*, *Normal*, *Overweight*, dan *Obese*. *World Health Organization* (WHO) mengklasifikasikan keempat kelas obesitas ini menggunakan *Body Mass Index* (BMI), yang didefinisikan sebagai

$$BMI = \frac{\text{berat}(kg)}{\text{tinggi}(m^2)}$$

Seseorang dengan $BMI < 18,5$ diklasifikasikan sebagai *underweight*, $18,5 \leq BMI < 25$ sebagai normal, $25 \leq BMI < 30$ sebagai *overweight*, dan $BMI \geq 30$ sebagai *obese* [10]. Adapun rincian mengenai tipe, deskripsi, dan distribusi semua variabel ditunjukkan pada Tabel 1.

2.1 Strategi Pemodelan

Pemodelan dilakukan dalam dua tahap. Tahap pertama melibatkan pelatihan tiga model *decision tree* secara terpisah, masing-masing menggunakan variabel dari masing-masing domain: Biologis, Pola Makan, atau Aktivitas Fisik. Pendekatan ini dipilih karena sifatnya yang mudah diinterpretasikan sekaligus memungkinkan perbandingan akurasi prediksi antar-domain. Kedua, dilakukan pemodelan XGBoost secara komprehensif, yaitu dengan melatih model XGBoost menggunakan semua 14 variabel prediktor. Pemodelan dilakukan dengan seed 123 untuk menjamin reproduksibilitas. XGBoost dipilih karena kemampuannya dalam menangani masalah klasifikasi dengan ketidakseimbangan data serta kompatibilitasnya dengan SHAP (*SHapley Additive Explanations*), yang memungkinkan interpretasi kontribusi variabel, baik secara global maupun lokal [11]. Penelitian sebelumnya juga menyarankan bahwa model berbasis pohon seperti

XGBoost, jika dikombinasikan dengan teknik *oversampling*, dapat meningkatkan akurasi klasifikasi pada dataset tidak seimbang [12], [13].

Data dibagi menjadi data latih (80%) dan data uji (20%) dengan *stratified sampling* untuk menjaga distribusi tiap kelas. Semua tahap praproses dan *tuning* model dilakukan hanya pada data latih guna mencegah *data leakage*. Tahap praproses meliputi:

- i. konversi variabel kategorial menjadi faktor,
- ii. *one-hot encoding* untuk variabel nominal maupun ordinal, dan
- iii. normalisasi variabel numerik. Adapun untuk mengatasi ketidakseimbangan kelas, digunakan SMOTENC *oversampling*, yang menghasilkan sampel sintetik pada kelas minoritas [14].

Setelah tahap praproses dan pembagian data, setiap model dievaluasi secara konsisten dengan metrik yang sama untuk memastikan perbandingan yang adil antar-pendekatan. Pendekatan ini tidak hanya menyoroti performa tiap domain secara individual, tetapi juga memberikan dasar untuk menilai kontribusi relatif setiap kelompok variabel terhadap klasifikasi obesitas secara keseluruhan. Tahap berikutnya menjelaskan secara lebih rinci penerapan model XGBoost beserta parameter utama yang dioptimasi selama proses pelatihan.

2.2 XGBoost (*eXtreme Gradient Boosting*)

XGBoost merupakan implementasi dari algoritma *gradient boosting* yang diusulkan oleh [15]. XGBoost adalah algoritma *ensemble learning* berbasis *decision tree* yang menggabungkan banyak *decision tree* sebagai *weak learners* untuk membentuk model prediksi yang akurat. Prinsip utamanya adalah membangun sebuah model *ensemble* secara berurutan, dimana *decision tree* berikutnya mengoreksi galat residual dari *decision tree* sebelumnya. Model ini menggunakan sebanyak K fungsi aditif untuk memprediksi y , yaitu sebagaimana ditunjukkan oleh persamaan matematis berikut.

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F}$$

dengan \mathcal{F} menyatakan himpunan *decision tree* sehingga f_k merujuk pada sebuah *decision tree* independen ke- k , x_i adalah peubah input, dan \hat{y}_i adalah prediksi dari peubah target y_i . XGBoost meminimalkan fungsi objektif berikut ini [16].

$$\mathcal{L}(\phi) = \sum_i l(y_i, \hat{y}_i) + \sum_k \Omega(f_k)$$

$l(y_i, \hat{y}_i)$ adalah *loss function* atau galat prediksi antara nilai prediksi \hat{y}_i dan nilai target y_i . Adapun $\Omega(f_k)$ merupakan *regularization term* yang mengontrol kompleksitas model dan didefinisikan sebagai

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$$

dengan T adalah jumlah daun dalam *decision tree*, γ adalah parameter regulasi yang menambahkan penalti ketika menambahkan daun, λ adalah parameter regulasi L2 yang mengontrol bobot daun w . Secara keseluruhan, penggunaan $\Omega(f_k)$ membantu model untuk menghindari *overfitting*.

Tabel 1. Jenis, Nilai, dan Deskripsi Variabel

| Variabel | Tipe | Nilai/Unit | Deskripsi |
|------------------------|---------|---|---|
| Biologis | | | |
| • Jenis Kelamin | Nominal | Perempuan (898), Laki-Laki (712) | Mengindikasikan jenis kelamin responden |
| • Usia | Rasio | 18 – 54 tahun | Merepresentasikan usia responden dalam satuan tahun |
| • Tinggi Badan | Rasio | 150 – 193 | Mengukur tinggi responden dalam satuan centimeter |
| • Riwayat Keluarga | Nominal | Ya (266), Tidak (1344) | Menunjukkan ada tidaknya riwayat kelebihan berat badan (<i>overweight</i>) atau obesitas (<i>obese</i>) pada anggota keluarga |
| Pola Makan | | | |
| • Makanan Cepat Saji | Nominal | Ya (436), Tidak (1174) | Mengindikasikan apakah responden mengonsumsi makanan cepat saji atau tidak |
| • Frekuensi Sayuran | Ordinal | Sangat Jarang (400), Sesekali (708), Selalu (502) | Merefleksikan seberapa sering responden mengonsumsi sayuran |
| • Jumlah Makan Berat | Ordinal | 1 – 2 kali sehari (444), 3 kali sehari (928), 3+ kali sehari (238) | Mewakili seberapa sering responden makan dalam sehari (makanan utama/berat) |
| • Camilan | Ordinal | Sangat Jarang (346), Sesekali (564), Sering (417), Selalu (283) | Mendeskrripsikan seberapa sering responden mengonsumsi camilan antar makanan utama/berat |
| • Konsumsi Air | Ordinal | <1 L (456), 1 – 2 L (523), >2 L (631) | Menghitung berapa liter air yang dikonsumsi responden dalam sehari |
| • Pemantauan Kalori | Nominal | Ya (286), Tidak (1324) | Menunjukkan apakah responden melakukan pelacakan terhadap kalori yang masuk atau tidak |
| Aktivitas Fisik | | | |
| • Merokok | Nominal | Ya (492), Tidak (118) | Mengindikasikan apakah responden merokok atau tidak |
| • Olahraga | Ordinal | 0 hari (206), 1–2 hari (290), 3–4 hari (370), 5–6 hari (358), 6+ hari (386) | Mengukur seberapa sering responden berolahraga dalam sepekan |
| • Durasi Layar | Ordinal | 0–2 jam (382), 3–5 jam (826), >5 jam (402) | Mengukur berapa lama waktu yang dihabiskan responden di depan layar, mewakili seberapa aktif responden sehari-hari |

| Variabel | Tipe | Nilai/Unit | Deskripsi |
|--------------------------|---------|---|--|
| • Transportasi | Nominal | Mobil (660), Sepeda Motor (94), Sepeda (116), Transportasi Publik (602), Jalan Kaki (138) | Mengidentifikasi mode transportasi utama yang digunakan oleh responden |
| Variabel Dependen | | | |
| • Kelas | Nominal | <i>Underweight</i> (73), <i>Normal</i> (658), <i>Overweight</i> (592), <i>Obese</i> (287) | Mengategorikan status berat masing-masing responden |

2.3 Evaluasi Kinerja Model

Kinerja model dievaluasi menggunakan beberapa metrik klasifikasi berikut [17], [18].

$$Precision_{macro} = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FP_i}$$

$$Recall_{macro} = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FN_i}$$

$$Specificity_{macro} = \frac{1}{K} \sum_{i=1}^K \frac{TN_i}{TN_i + FP_i}$$

$$F1_{macro} = \frac{1}{K} \sum_{i=1}^K \frac{2 \cdot Precision_i \cdot Recall_i}{Precision_i + Recall_i}$$

$$Balanced Accuracy_{macro} = \frac{1}{K} \sum_{i=1}^K \frac{Recall_i + Specificity_i}{2}$$

dengan TP_i, FP_i, TN_i, FN_i masing-masing merujuk pada *true positive, false positive, true negative, dan false negative*, sementara $i = 1,2,3,4$ merujuk pada kelas ke- i dari $K = 4$ kelas yang dievaluasi. Kemudian, karena variabel dependen terdiri atas empat kelas, maka metrik-metrik tersebut perlu diagregatkan. Pada penelitian ini, kami menggunakan *macro average*.

Selain itu, penelitian ini juga menggunakan kurva ROC-AUC dengan metode Hand-Till dimana i dan j adalah kelas [19], [20].

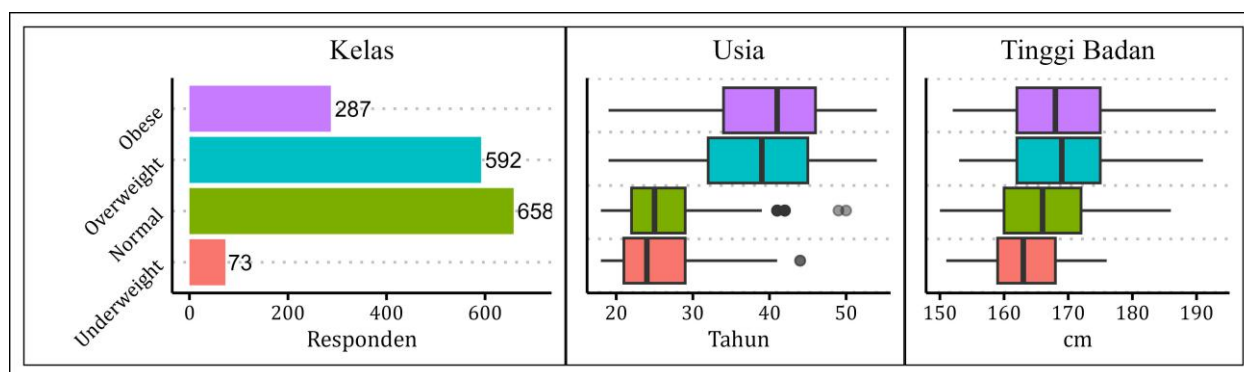
$$AUC_{Hand-Till} = \frac{2}{K(K-1)} \sum_{i < j} \frac{AUC_{ij} + AUC_{ji}}{2}$$

3 Hasil dan Pembahasan

3.1 Eksplorasi Data

Pada tahap eksplorasi awal, ditemukan bahwa data yang digunakan bersih tanpa data hilang pada variabel manapun. Variabel terdiri dari dua variabel kontinu, yaitu usia dan tinggi badan, sedangkan variabel lainnya bersifat kategorial, baik ordinal maupun nominal. Variabel-variabel kategorial tersebut awalnya dinyatakan sebagai bilangan bulat, kemudian diubah menjadi *factor* sebelum proses pemodelan. Kemudian, karena XGBoost sensitif terhadap skala variabel dalam optimisasi matriks internalnya, maka dilakukan pula praproses seperti *one-hot encoding* pada variabel nominal serta normalisasi pada variabel numerik.

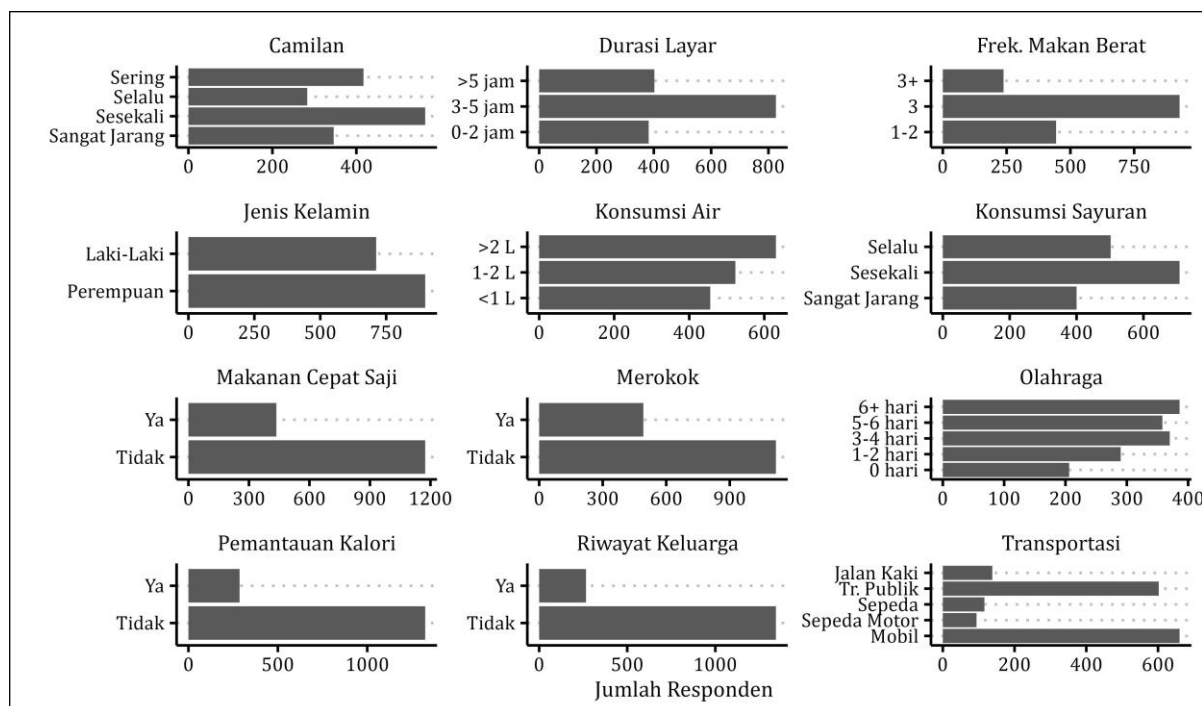
Untuk memahami karakteristik data, digunakan *boxplot* dan *bar plot* untuk menggambarkan distribusi setiap variabel. Gambar 1 menampilkan distribusi variabel numerik berdasarkan setiap kelas obesitas. Plot tersebut juga menunjukkan jumlah responden untuk setiap kelas obesitas.



Gambar 1. Distribusi Tingkat Obesitas (Kelas) dan Variabel Numerik

Seperti terlihat pada Gambar 1, data mengalami ketidakseimbangan kelas (*class imbalance*) karena hanya terdapat 73 responden *underweight* dan 287 responden *obese*, masing-masing sekitar 11% dan 44% dari jumlah responden kelas *normal*. Ketidakseimbangan yang cukup besar ini dapat memengaruhi akurasi klasifikasi, sehingga diperlukan penanganan untuk meningkatkan akurasi.

Gambar 1 juga menunjukkan distribusi usia dan tinggi badan. Terlihat jelas bahwa responden *overweight* dan *obese* cenderung berusia lebih tua dibandingkan responden pada kelas *normal* dan *underweight*. Hal ini mengindikasikan bahwa usia mungkin menjadi variabel yang baik untuk tingkat obesitas, sedangkan tinggi badan mungkin bukan variabel yang sebagus usia. Sementara itu, jumlah responden untuk setiap kelas pada variabel kategorial ditunjukkan pada Gambar 2.



Gambar 2. Distribusi Variabel Kategorial

Seperti ditunjukkan pada Gambar 2, selain jenis kelamin dan konsumsi air, proporsi variabel lainnya relatif tidak seimbang. Namun, uji chi-square pada variabel kategorial menunjukkan adanya asosiasi yang kuat dengan variabel dependen, sebagaimana ditampilkan pada Tabel 2.

Tabel 2. Hasil Uji Chi-Square untuk Asosiasi antara Variabel Kategorial dan Kelas Obesitas

| Variabel | χ^2 | df | <i>p-value</i> |
|-----------------------|----------|----|----------------|
| Jenis Kelamin | 156,1730 | 3 | 1,23E-33 |
| Riwayat Keluarga | 180,3889 | 3 | 7,27E-39 |
| Makanan Cepat Saji | 250,0021 | 3 | 6,54E-54 |
| Konsumsi Sayuran | 554,6705 | 6 | 1,4E-116 |
| Frekuensi Makan Berat | 576,6152 | 6 | 2,6E-121 |
| Camilan | 86,7743 | 9 | 7,21E-15 |
| Konsumsi Air | 72,5538 | 6 | 1,22E-13 |
| Pemantauan Kalori | 156,1901 | 3 | 1,22E-33 |
| Merokok | 185,4534 | 3 | 5,86E-40 |
| Olahraga | 409,9477 | 12 | 2,96E-80 |
| Durasi Layar | 46,3502 | 6 | 2,52E-08 |

3.2 Pemodelan Berbasis Domain Menggunakan *Decision Tree*

Model *decision tree* dapat terpengaruh oleh ketidakseimbangan kelas. Untuk mengatasi masalah ini, digunakan pembagian data latih dan uji menggunakan stratifikasi dengan rasio 80:20. Penggunaan teknik *oversampling* juga dipertimbangkan sebelum memilih model terbaik.

Sebelum menerapkan XGBoost pada seluruh data latih, *decision tree* digunakan untuk membentuk tiga model berdasarkan domain, yaitu: model yang dilatih hanya dengan variabel domain biologis, model dengan variabel domain pola makan, dan model dengan variabel domain aktivitas fisik. Hasil dari pemodelan ini digunakan sebagai dasar pembandingan terhadap hasil XGBoost. Selain itu, *decision tree* mudah diinterpretasikan sehingga hasilnya dapat digunakan untuk menguji konsistensi interpretasi model XGBoost.

Penyetelan (*tuning*) parameter dilakukan pada tiga *hyperparameter*, yaitu *cost complexity* pada rentang 0 hingga 0,1; *tree depth* pada rentang 2 hingga 5 untuk menjaga kemudahan interpretasi; serta *min n* pada rentang 2 hingga 40.

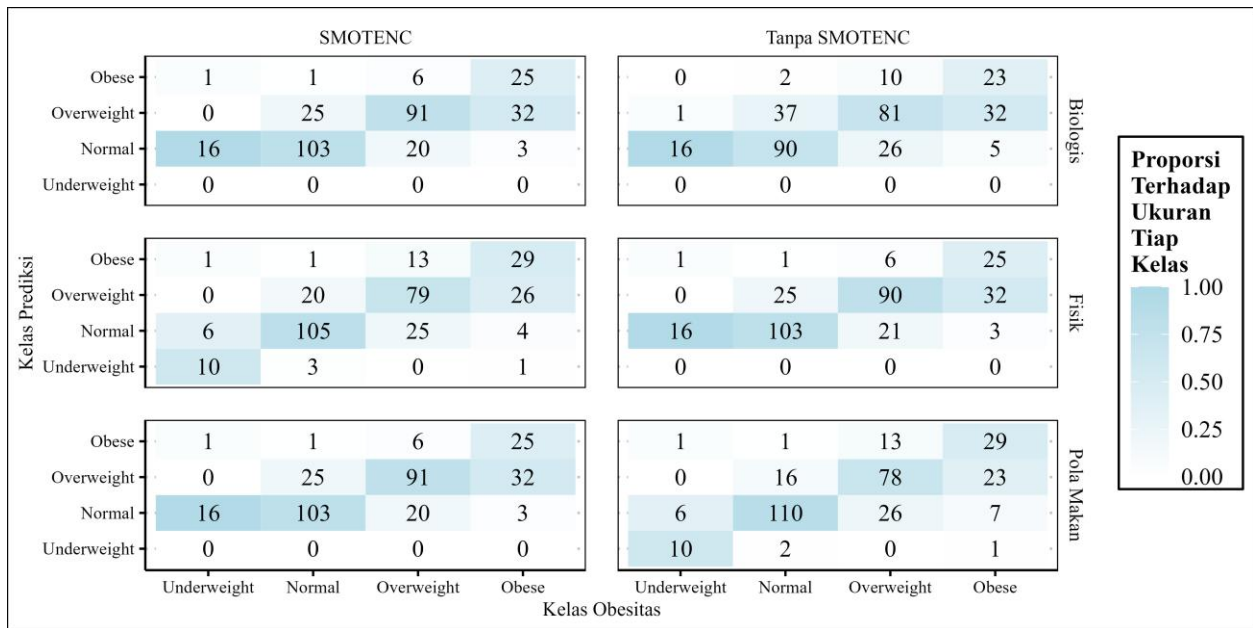
Proses penyetelan ini di-*resample* menggunakan *10-fold cross validation* yang distratifikasi terhadap variabel dependen. Hasil evaluasi menggunakan metrik *balanced accuracy* dan ROC-AUC. Selain itu, karena terdapat ketidakseimbangan kelas yang cukup parah, teknik SMOTENC digunakan untuk menyeimbangkan jumlah responden pada kelas *underweight* dan *obese* agar mendekati jumlah pada kelas *normal*.

Kami melakukan pengujian terhadap 128 kombinasi *hyperparameter* dan praproses dengan evaluasi terhadap total sebanyak 384 model. Tabel 3 menampilkan konfigurasi *hyperparameter* optimal untuk setiap domain dan metrik.

Tabel 3. Kinerja Model *Decision Tree* Berbasis Domain Terbaik Berdasarkan Hasil Validasi

| Metrik | Domain | SMOTENC | Rata-Rata | SE | Cost Compl. | Tree Depth | Min n |
|--------------------------|------------|---------|-----------|----------|-------------|------------|-------|
| <i>Balanced Accuracy</i> | Biologis | Tidak | 0,658815 | 0,007867 | 1E-10 | 2 | 2 |
| | | Ya | 0,646590 | 0,010797 | 1E-10 | 5 | 2 |
| | Pola Makan | Tidak | 0,672292 | 0,004816 | 1E-10 | 5 | 27 |
| | | Ya | 0,683967 | 0,019066 | 1E-10 | 4 | 2 |
| | Fisik | Tidak | 0,667620 | 0,013189 | 1E-10 | 4 | 40 |
| | | Ya | 0,681486 | 0,012921 | 1E-10 | 5 | 40 |
| ROC-AUC | Biologis | Tidak | 0,752197 | 0,012748 | 1E-10 | 5 | 40 |
| | | Ya | 0,766580 | 0,011036 | 1E-10 | 5 | 2 |
| | Pola Makan | Tidak | 0,781235 | 0,010594 | 1E-10 | 5 | 14 |
| | | Ya | 0,771158 | 0,013804 | 1E-10 | 5 | 40 |
| | Fisik | Tidak | 0,762197 | 0,011143 | 1E-10 | 5 | 27 |
| | | Ya | 0,771984 | 0,006440 | 1E-10 | 3 | 40 |

Secara umum, penggunaan SMOTENC meningkatkan nilai ROC-AUC untuk domain Biologis dan Fisik, meskipun peningkatan pada *balanced accuracy* relatif kecil. Menariknya, pada domain Pola Makan, model tanpa SMOTENC menghasilkan *balanced accuracy* yang lebih baik, dengan penurunan ROC-AUC yang hanya sedikit.

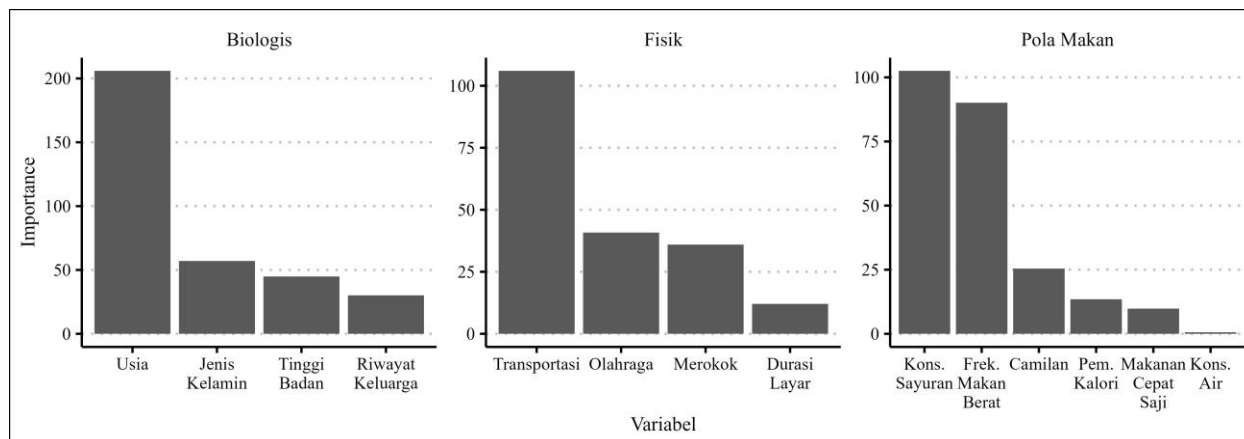


Gambar 3. *Confusion Matrix* Model Terbaik pada Data Uji

Untuk memahami implikasi dari penggunaan SMOTENC, dilakukan evaluasi lebih lanjut terhadap hasil prediksi model terbaik menggunakan *confusion matrix*, sebagaimana ditunjukkan pada Gambar 3. Warna pada matriks disesuaikan dengan jumlah total responden pada setiap kelas.

Gambar 3 menggambarkan kinerja prediksi pada masing-masing kelas untuk setiap domain, baik dengan maupun tanpa SMOTENC. Pada domain Biologis dan Fisik, penggunaan SMOTENC secara signifikan mengurangi kesalahan klasifikasi pada kelas minoritas *Underweight* dan *Obese* yang umumnya kurang terwakili. Sebaliknya, pada domain Pola Makan, model tanpa SMOTENC justru menunjukkan kinerja keseluruhan yang lebih baik, khususnya dalam mengklasifikasikan individu *Underweight* secara benar, mengindikasikan bahwa SMOTENC dalam konteks ini mungkin menambahkan *noise* atau menyebabkan *overfitting*.

Berdasarkan hasil ini, dipilih model dengan SMOTENC sebagai model terbaik pada domain Biologis dan Fisik, serta model tanpa SMOTENC sebagai model terbaik pada domain Pola Makan. *Variable importance* dari ketiga konfigurasi optimal ini ditunjukkan pada Gambar 4 untuk memberikan interpretasi lebih lanjut.



Gambar 4. *Variable Importance* pada Model Optimal untuk Setiap Domain

Gambar 4 menunjukkan skor *variable importance* dari model terpilih pada setiap domain. Pada domain biologis, usia merupakan variabel paling berpengaruh, diikuti oleh tinggi badan, sedangkan jenis kelamin dan riwayat keluarga memiliki kontribusi lebih kecil. Pada domain pola makan, frekuensi konsumsi sayur, dan jumlah makan utama merupakan variabel terpenting, sementara variabel seperti konsumsi makanan cepat saji, kebiasaan makan camilan, pemantauan kalori, dan konsumsi air memiliki peran yang lebih kecil. Pada domain fisik, jenis transportasi menempati posisi tertinggi, diikuti oleh frekuensi olahraga, kebiasaan merokok, dan durasi layar.

3.3 Pemodelan Lengkap Menggunakan XGBoost

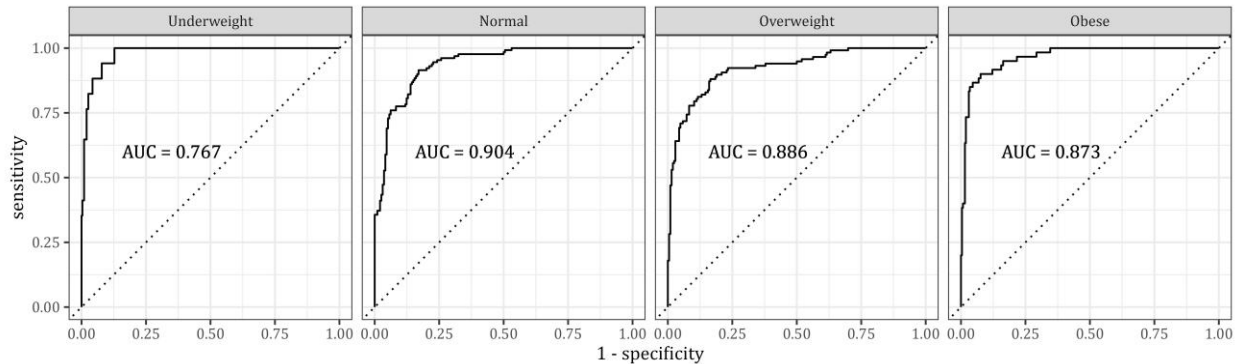
Setelah melakukan pemodelan *decision tree* berbasis domain, selanjutnya dilakukan klasifikasi penuh menggunakan XGBoost. Berbeda dengan model berbasis domain, model ini memanfaatkan semua peubah secara bersamaan, sehingga memungkinkan model menangkap potensi interaksi antar-domain dan meningkatkan kinerja klasifikasi secara keseluruhan.

Untuk memastikan model ter-*tuning* dengan baik, dilakukan optimasi *hyperparameter* dengan memvariasikan jumlah variabel yang dipilih secara acak pada setiap pemisahan (*mtry*), kedalaman pohon (*tree depth*), laju pembelajaran (*learning rate*), pengurangan kerugian minimum (*gamma*), dan ukuran sampel (*subsample size*). Lima konfigurasi dengan kinerja terbaik ditampilkan pada Tabel 4.

Tabel 4. Kinerja Lima Model XGBoost Terbaik Berdasarkan Hasil Validasi

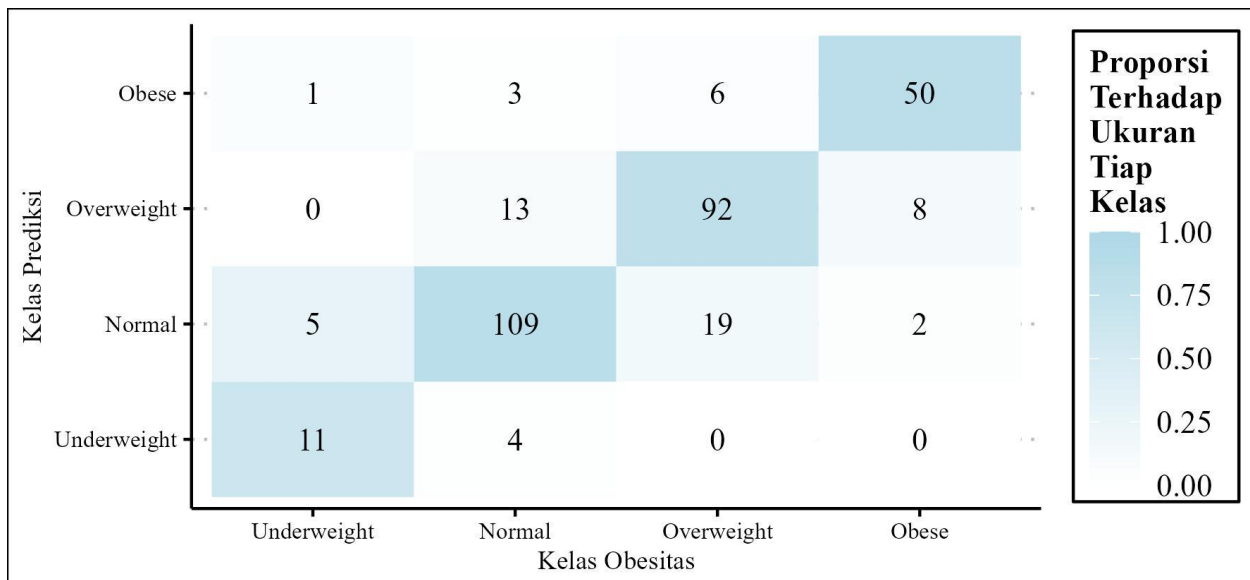
| <i>mtry</i> | <i>Tree Depth</i> | <i>Learning Rate</i> | <i>Loss Reduction</i> | <i>Sample Size</i> | <i>Balanced Accuracy</i> | | <i>ROC-AUC</i> | |
|-------------|-------------------|----------------------|-----------------------|--------------------|--------------------------|----------|----------------|----------|
| | | | | | Rata-Rata | SE | Rata-Rata | SE |
| 5 | 6 | 0,042793 | 0,000425 | 0,543974 | 0,858309 | 0,012067 | 0,958520 | 0,004538 |
| 8 | 12 | 0,009248 | 5,87E-09 | 0,372300 | 0,851580 | 0,012530 | 0,954712 | 0,004702 |
| 11 | 7 | 0,004013 | 2,92E-08 | 0,440170 | 0,835938 | 0,017195 | 0,946605 | 0,006345 |
| 1 | 14 | 0,015457 | 9,26E-10 | 0,915992 | 0,823735 | 0,015496 | 0,943914 | 0,004852 |
| 8 | 13 | 6,59E-08 | 0,608131 | 0,730982 | 0,814328 | 0,016018 | 0,931375 | 0,007393 |

Dari kombinasi konfigurasi yang diuji, model dengan $mtry = 5$, $tree\ depth = 6$, $learning\ rate = 0,043$, $minimum\ loss\ reduction = 0,000425$, dan $sample\ size = 0,544$ memberikan kinerja terbaik, dengan *balanced accuracy* sebesar 0,858 dan ROC-AUC sebesar 0,959. Nilai ini lebih tinggi dibandingkan kinerja ketiga model berbasis domain, yang menunjukkan bahwa penggabungan fitur antar-domain memberikan nilai prediksi yang signifikan.



Gambar 5. ROC-AUC untuk Setiap Kelas Obesitas

Kinerja per kelas dari model terbaik divisualisasikan pada Gambar 5, yang menampilkan kurva ROC-AUC untuk masing-masing kelas. Kurva tersebut mengonfirmasi kemampuan model yang kuat dalam membedakan keempat kategori obesitas, dengan nilai AUC tertinggi pada kelas *Normal* dan *Overweight*. Confusion matrix pada Gambar 6 menunjukkan bahwa kesalahan klasifikasi paling sering terjadi antara kelas *Obese* dan *Overweight*.



Gambar 6. Matriks Kebingungan Model XGBoost Optimal

Klasifikasi kelas *Underweight* dan *Obese* menjadi bagian yang paling menantang bagi model karena adanya ketidakseimbangan kelas. Namun, penggunaan XGBoost dengan SMOTENC telah meningkatkan akurasi prediksi pada kedua kelas ini, terlihat dari jumlah kasus yang teridentifikasi

dengan benar yang meningkat menjadi 11 dan 50 kasus, dibandingkan sebelumnya 10 dan 29 kasus pada *decision tree* berbasis domain seperti yang ditunjukkan pada Gambar 3.

Untuk memahami kinerja pada tingkat kelas, Tabel 5 menampilkan metrik evaluasi utama pada data uji untuk setiap kelas. Berdasarkan Tabel 5, nilai *recall* untuk kelas *Underweight* tetap relatif rendah (0,183), menunjukkan bahwa kelas ini masih sulit diprediksi dengan benar, bahkan pada model yang menggunakan keseluruhan variabel. Meski demikian, model berhasil mencapai nilai *recall* dan *precision* yang tinggi untuk kelas *Normal* dan *Overweight*, dengan *F1-score* di atas 0,84 pada keduanya.

Hasil ini menunjukkan bahwa meskipun XGBoost secara signifikan meningkatkan kinerja klasifikasi, terutama pada kelas mayoritas dan kelas menengah, masih terdapat tantangan dalam mengidentifikasi kasus minoritas seperti *Underweight*. Walau demikian, model XGBoost ini menjadi kandidat terbaik untuk interpretasi lebih lanjut menggunakan nilai SHAP.

Tabel 5. Kinerja Model XGBoost dan *Decision Tree* Berbasis Domain pada Data Uji per Kelas

| <i>Class</i> | <i>Model</i> | <i>Precision</i> | <i>Recall</i> | <i>Specificity</i> | <i>F1</i> | <i>Balanced Accuracy</i> |
|--------------------|---------------|------------------|---------------|--------------------|-----------|--------------------------|
| <i>Underweight</i> | XGBoost | 0,733333 | 0,183333 | 0,984436 | 0,293333 | 0,583885 |
| | DT Biologis | NaN | 0,000000 | 1,000000 | NaN | 0,500000 |
| | DT Pola Makan | 0,769231 | 0,104167 | 0,986364 | 0,183486 | 0,545265 |
| | DT Fisik | 0,714286 | 0,10101 | 0,981567 | 0,176991 | 0,541288 |
| <i>Normal</i> | XGBoost | 0,807407 | 0,893443 | 0,856354 | 0,848249 | 0,874898 |
| | DT Biologis | 0,725352 | 0,725352 | 0,748387 | 0,725352 | 0,736870 |
| | DT Pola Makan | 0,738255 | 0,743243 | 0,750000 | 0,740741 | 0,746622 |
| | DT Fisik | 0,750000 | 0,719178 | 0,771242 | 0,734266 | 0,745210 |
| <i>Overweight</i> | XGBoost | 0,830357 | 0,853211 | 0,900000 | 0,841629 | 0,876606 |
| | DT Biologis | 0,614865 | 0,812500 | 0,691892 | 0,700000 | 0,752196 |
| | DT Pola Makan | 0,666667 | 0,812500 | 0,792553 | 0,732394 | 0,802527 |
| | DT Fisik | 0,632000 | 0,831579 | 0,757895 | 0,718182 | 0,794737 |
| <i>Obese</i> | XGBoost | 0,836066 | 0,560440 | 0,955157 | 0,671053 | 0,757798 |
| | DT Biologis | 0,757576 | 0,290698 | 0,960396 | 0,420168 | 0,625547 |
| | DT Pola Makan | 0,659091 | 0,367089 | 0,929577 | 0,471545 | 0,648333 |
| | DT Fisik | 0,659091 | 0,349398 | 0,928230 | 0,456693 | 0,638814 |

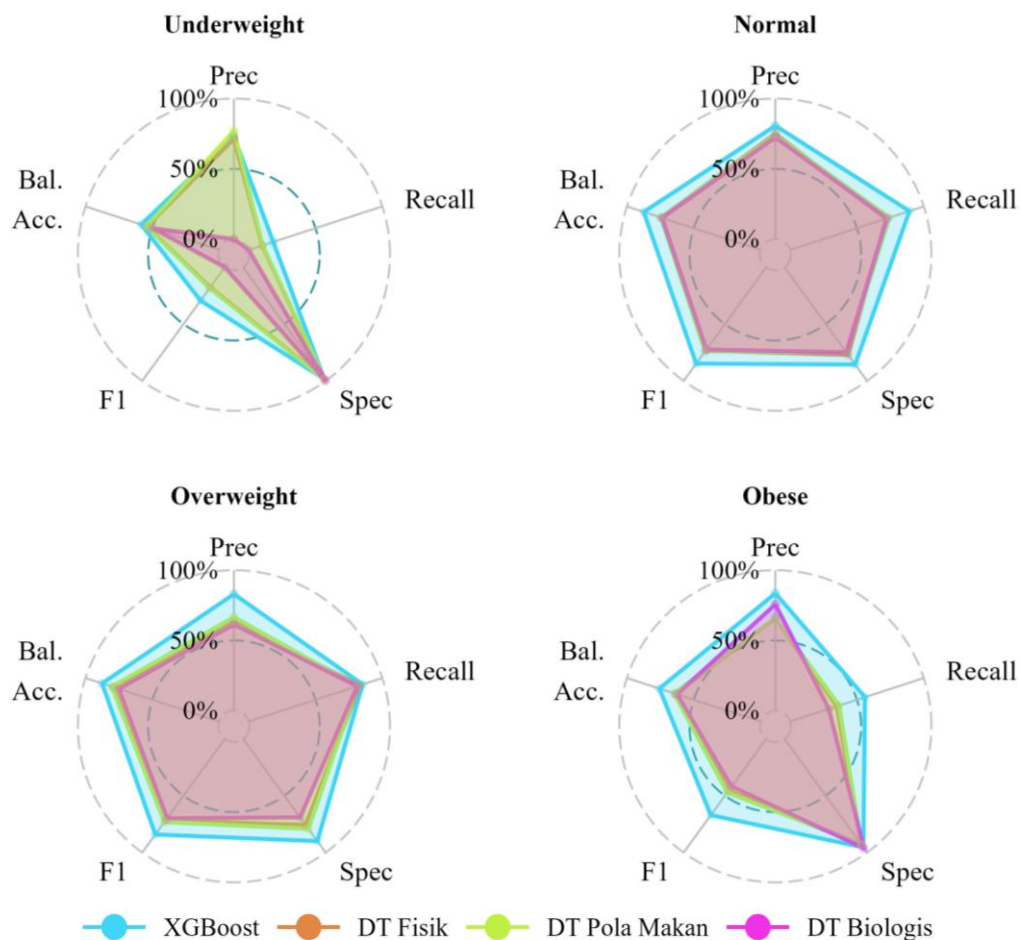
3.4 Perbandingan Model dan Interpretasi

Untuk menilai kinerja relatif dan kemudahan interpretasi dari setiap pendekatan pemodelan, dilakukan perbandingan antara model XGBoost dan tiga model *decision tree* berbasis domain menggunakan lima metrik evaluasi. Hasil perbandingan ini divisualisasikan pada Gambar 7, yang

menampilkan *spider plot* untuk *precision*, *recall*, *specificity*, *F1-score*, dan *balanced accuracy* pada setiap kelas obesitas.

Gambar 7 menunjukkan secara jelas bahwa XGBoost memiliki kinerja yang lebih unggul dan konsisten di semua kelas. Area yang tertutup oleh garis XGBoost pada *radar plot* adalah yang terbesar, menandakan kekuatan pada semua metrik evaluasi. Sebaliknya, model *decision tree* menunjukkan variabilitas yang lebih tinggi dan cakupan yang lebih terbatas, terutama pada kelas *Underweight* dan *Obese*, yang diketahui rentan terhadap ketidakseimbangan kelas dan interaksi variabel yang kompleks.

Di antara tiga model berbasis domain, domain Pola Makan memiliki kinerja terbaik secara keseluruhan, dengan skor metrik yang relatif stabil. Temuan ini konsisten dengan hasil sebelumnya yang menunjukkan bahwa variabel Pola Makan, seperti frekuensi konsumsi sayur dan makan berat, penting untuk membedakan kelas obesitas menengah seperti *Normal* dan *Overweight*.



Gambar 7. Ringkasan Prediksi Model Optimal pada Data Uji

Selain itu, temuan lain yang menarik dari pemodelan berbasis domain adalah seberapa buruknya DT Biologis dalam klasifikasi, utamanya dalam mengklasifikasikan kelas *underweight*. Sebagaimana dapat dilihat secara lebih detail pada Tabel 5, *precision* dan F1-score dari model DT

Biologis pada kelas *underweight* berupa NaN (*Not a Number*). Hal ini terjadi karena tidak ada amatan yang diklasifikasikan sebagai *underweight* oleh DT Biologis. Akibatnya, *FP* dan *TP* bernilai 0, sehingga

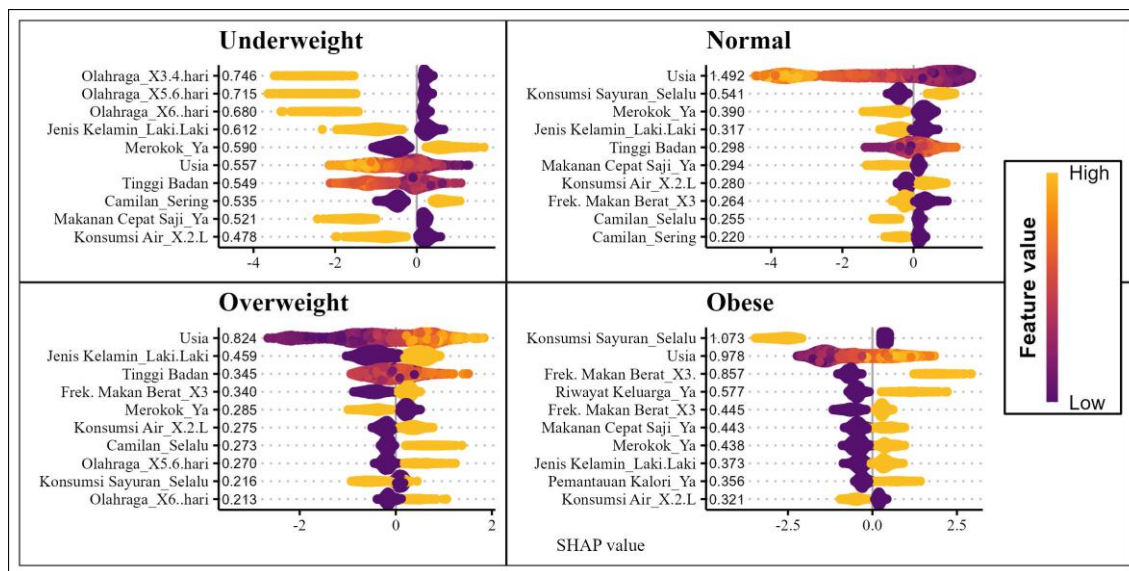
$$precision = \frac{TP}{TP + FP} = \frac{0}{0 + 0} = NaN$$

Kemudian, karena

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

Maka nilai *F1-score* juga menjadi NaN. Kedua hasil ini menunjukkan bahwa domain biologis tidak cukup untuk memprediksi kelas *underweight*, terutama jika menggunakan model *Decision Tree*.

Jika merujuk pada Gambar 7, kelas *underweight* sebenarnya memang merupakan kelas yang paling sulit untuk diprediksi, meskipun telah menggunakan XGBoost sekalipun. Meskipun begitu, penggunaan XGBoost setidaknya masih menghasilkan hasil yang lebih baik dibandingkan ketiga model DT. Untuk memahami alasan di balik keunggulan XGBoost, dilakukan analisis terhadap logika internal model menggunakan SHAP sebagaimana ditunjukkan pada Gambar 8.



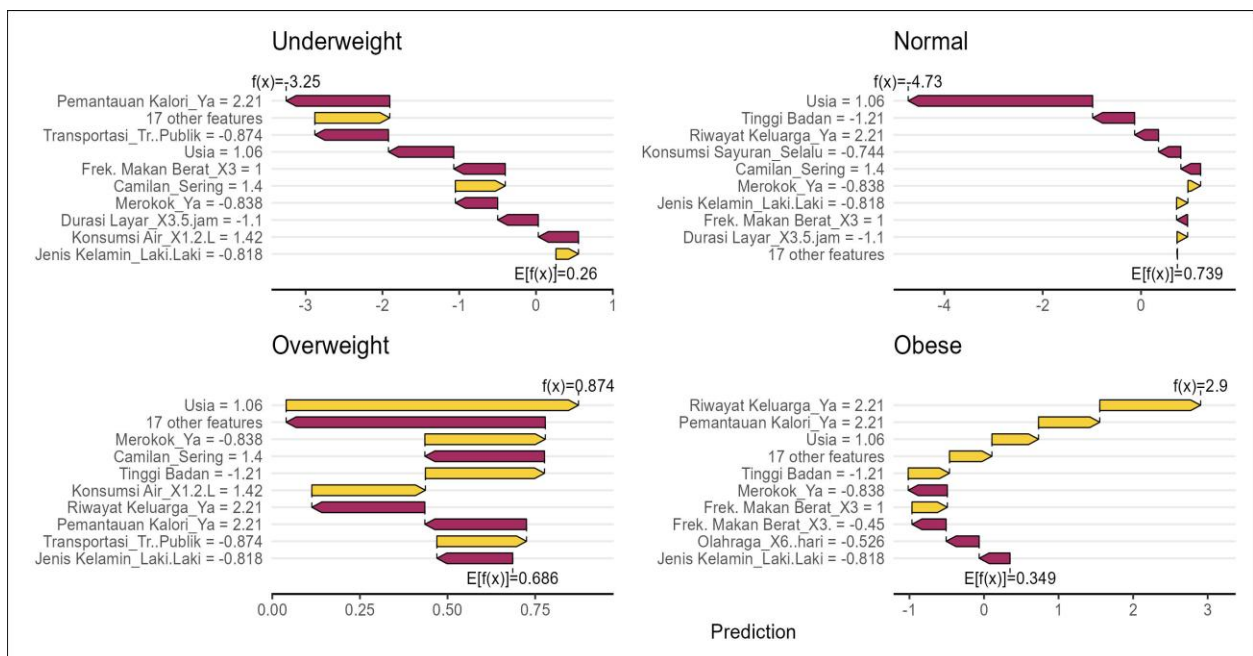
Gambar 8. SHAP Summary Beeswarm dari Model XGBoost Optimal

Gambar 8 menampilkan *SHAP summary plot* yang memvisualisasikan besaran dan arah kontribusi masing-masing variabel terhadap semua hasil prediksi. Variabel paling penting adalah Usia yang memengaruhi prediksi ke arah kedua ekstrem. Dalam hal ini, usia yang lebih tua cenderung mendorong model XGBoost untuk mengklasifikasikan responden sebagai *Overweight* dan *Obese*. Fitur penting lainnya termasuk frekuensi konsumsi sayur, frekuensi makan berat, jenis transportasi, dan kebiasaan makan camilan.

Menariknya, variabel-variabel dengan kontribusi tertinggi berdasarkan plot SHAP pada Gambar 8 mencakup semua domain, menguatkan interpretasi bahwa model berbasis domain hanya menangkap sebagian kecil dari keragaman data. Sebagai contoh:

- Usia (Biologis) adalah variabel paling dominan secara keseluruhan.
- Konsumsi sayur dan frekuensi makan berat (Pola Makan) juga sangat berpengaruh.
- Jenis transportasi dan aktivitas fisik (Fisik) turut memberikan kontribusi yang berarti, meskipun efeknya lebih lemah dan bergantung pada interaksi dengan variabel lain.

Temuan dari SHAP ini melengkapi hasil *variable importance* pada *decision tree* dan memberikan pandangan lebih dalam terkait pentingnya variabel pada tingkat domain. Hal ini juga memperkuat justifikasi penggunaan model seperti XGBoost, yang mampu menggabungkan pola data yang lemah hingga sedang dari berbagai domain menjadi prediksi yang akurat.



Gambar 9. Contoh prediksi akurat oleh XGBoost (ID = 63)

yang tidak terdeteksi oleh semua model *Decision Tree*

Sebagai perbandingan, Gambar 9 menampilkan satu contoh kasus dari data uji (ID = 63). Dalam contoh ini, subjek diklasifikasikan dengan benar sebagai *Obese* oleh XGBoost, sementara ketiga *decision tree* berbasis domain salah mengklasifikasikannya sebagai *Overweight*. Perbedaan ini mengilustrasikan keterbatasan ketiga model *decision tree* yang terbatas pada domain masing-masing: Ketiga model tersebut dapat melewatkan interaksi atau efek gabungan dari variabel lintas domain. Sebaliknya, XGBoost mampu mengintegrasikan interaksi tersebut sehingga menghasilkan prediksi yang lebih akurat, bahkan pada kasus yang sulit seperti kasus ID = 63.

Secara keseluruhan, bagian ini menunjukkan bahwa meskipun model berbasis domain memberikan wawasan terkait perilaku kelompok variabel, model XGBoost menawarkan kinerja prediktif yang lebih akurat sekaligus kerangka interpretasi yang lebih mendalam jika dilengkapi dengan SHAP.

3.5 Diskusi

Penelitian ini bertujuan untuk mengklasifikasikan tingkat obesitas menggunakan 14 variabel yang berasal dari domain biologis, pola makan, dan aktivitas fisik. Dengan menerapkan model *decision tree* berbasis domain yang interpretatif serta model XGBoost yang berperforma tinggi, penelitian ini mengeksplorasi akurasi prediktif sekaligus kemudahan interpretasi dari masing-masing pendekatan.

Hasil penelitian menunjukkan bahwa model XGBoost secara signifikan mengungguli *decision tree* berbasis domain pada semua metrik kinerja, menegaskan pentingnya integrasi semua variabel. Meskipun *decision tree* menawarkan kemudahan interpretasi, model ini tidak mampu menangkap pola kompleks yang dibutuhkan untuk klasifikasi obesitas secara akurat.

Di antara model berbasis domain, domain Pola Makan menunjukkan akurasi yang relatif lebih baik. Temuan ini selaras dengan literatur sebelumnya yang menekankan peran penting kebiasaan makan dalam hasil obesitas [6], [21]. Namun, analisis SHAP menunjukkan bahwa variabel yang penting dalam prediksi tersebar di ketiga domain, khususnya usia (biologis), konsumsi sayur dan frekuensi makan berat (pola makan), serta jenis transportasi (fisik). Hal ini menegaskan pentingnya interaksi lintas domain yang hanya dapat ditangkap secara efektif oleh model komprehensif seperti XGBoost [2], [22], [23], [24], [25].

Contoh prediksi pada responden dengan ID 63 lebih lanjut memperjelas temuan ini: Seluruh model berbasis domain salah mengklasifikasikan responden tersebut sebagai *Overweight*, sementara XGBoost berhasil mengidentifikasi tingkat obesitas dengan benar. Hal ini menyoroti keterbatasan dalam mengandalkan logika domain tunggal secara eksklusif.

Meskipun penerapan SMOTENC dan penyetelan *hyperparameter* membantu mengatasi ketidakseimbangan kelas dan mengoptimalkan kinerja model, masih terdapat keterbatasan dengan pendekatan SMOTENC. Model masih kesulitan dalam mengklasifikasikan kelompok *Underweight* karena representasi yang rendah pada data. Kondisi ini mengindikasikan perlunya penambahan data lebih lanjut atau penerapan teknik sampling alternatif untuk meningkatkan deteksi kelas minoritas.

Walaupun SHAP memungkinkan kemudahan interpretasi pasca-model (*post hoc*), pendekatan ini tidak sepenuhnya menggantikan kejelasan intuitif dari model yang lebih sederhana seperti *decision tree*. Untuk sistem pendukung keputusan dalam konteks kesehatan masyarakat

atau klinis, keseimbangan antara kemudahan interpretasi dan kekuatan prediktif tetap menjadi tantangan utama.

Selain keterbatasan terkait ketidakseimbangan kelas, penelitian ini juga belum mencakup proses validasi eksternal. Model dikembangkan dan diuji menggunakan satu sumber data dengan periode dan lokasi yang sama, sehingga generalisasi hasil ke populasi yang lebih luas masih perlu dikaji lebih lanjut. Faktor-faktor seperti perbedaan demografi, kebiasaan makan, dan gaya hidup antarwilayah berpotensi memengaruhi kinerja model. Oleh karena itu, penelitian lanjutan disarankan untuk melakukan validasi eksternal menggunakan data dari lokasi atau periode berbeda guna memastikan robustitas dan stabilitas model dalam berbagai konteks populasi.

4 Simpulan

Penelitian ini meneliti klasifikasi tingkat obesitas menggunakan 14 variabel yang dikelompokkan ke dalam tiga domain: biologis, pola makan, dan aktivitas fisik. Dengan membandingkan model *decision tree* berbasis domain dan model XGBoost komprehensif, penelitian ini bertujuan untuk menilai kontribusi relatif dari masing-masing domain sekaligus mengevaluasi manfaat integrasi variabel multi-domain.

Hasil penelitian menunjukkan bahwa meskipun model berbasis domain menawarkan kemudahan interpretasi, akurasi yang dihasilkan masih lebih rendah dibandingkan model XGBoost yang mengintegrasikan keseluruhan data. Analisis SHAP mengungkapkan bahwa variabel paling berpengaruh seperti Usia, Konsumsi Sayur, Frekuensi Makan Berat, dan Jenis Transportasi tersebar di seluruh domain. Hal ini mendukung kesimpulan bahwa interaksi lintas domain sangat penting untuk klasifikasi obesitas yang akurat.

Meskipun kinerja model secara keseluruhan cukup akurat, tantangan tetap ada pada klasifikasi kelas minoritas, khususnya kelompok *Underweight*. Keterbatasan ini disebabkan oleh ketidakseimbangan kelas, sehingga menekankan perlunya penambahan data atau penggunaan strategi *sampling* alternatif untuk meningkatkan deteksi pada kelas dengan representasi rendah.

5 Daftar Pustaka

- [1] T. Basu, A. Selman, A. P. Reddy, and P. H. Reddy, "Current Status of Obesity: Protective Role of Catechins," *Antioxidants* 2023, Vol. 12, Page 474, vol. 12, no. 2, p. 474, Feb. 2023, doi: 10.3390/ANTIOX12020474.
- [2] X. Jin *et al.*, "Pathophysiology of obesity and its associated diseases," *Acta Pharm Sin B*, vol. 13, no. 6, pp. 2403–2424, Jun. 2023, doi: 10.1016/J.APSB.2023.01.012.

-
- [3] M. Janić, A. Janež, M. El-Tanani, and M. Rizzo, “Obesity: Recent Advances and Future Perspectives,” *Biomedicines* 2025, Vol. 13, Page 368, vol. 13, no. 2, p. 368, Feb. 2025, doi: 10.3390/BIOMEDICINES13020368.
- [4] M. Tariqujjaman *et al.*, “Determinants of Double Burden of Malnutrition Among School Children and Adolescents in Urban Dhaka: A Multi-Level Analyses,” *Front Public Health*, vol. 10, p. 926571, Jul. 2022, doi: 10.3389/FPUBH.2022.926571/BIBTEX.
- [5] World Health Organization, “Obesity and overweight.” Accessed: Jul. 27, 2025. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>
- [6] J. Du, S. Yang, Y. Zeng, C. Ye, X. Chang, and S. Wu, “Visualization obesity risk prediction system based on machine learning,” *Scientific Reports* 2024 14:1, vol. 14, no. 1, pp. 1–9, Sep. 2024, doi: 10.1038/s41598-024-73826-6.
- [7] Z. Helforoush and H. Sayyad, “Prediction and classification of obesity risk based on a hybrid metaheuristic machine learning approach,” *Front Big Data*, vol. 7, p. 1469981, Sep. 2024, doi: 10.3389/FDATA.2024.1469981/BIBTEX.
- [8] N. Koklu and S. A. Sulak, “Predicting risk of obesity and meal planning to reduce the obese in adulthood using artificial intelligence,” *Endocrine*, vol. 78, no. 3, pp. 458–469, Dec. 2022, doi: 10.1007/S12020-022-03215-4/METRICS.
- [9] Y. C. Lee *et al.*, “Using Machine Learning to Predict Obesity Based on Genome-Wide and Epigenome-Wide Gene–Gene and Gene–Diet Interactions,” *Front Genet*, vol. 12, p. 783845, Jan. 2022, doi: 10.3389/FGENE.2021.783845/BIBTEX.
- [10] World Health Organization, “Moderate and severe thinness, underweight, overweight and obesity,” Nutrition Landscape Information System (NLiS). Accessed: Nov. 02, 2025. [Online]. Available: <https://apps.who.int/nutrition/landscape/help.aspx?menu=0&helpid=392&lang=EN>
- [11] S. M. Lundberg, P. G. Allen, and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” *Adv Neural Inf Process Syst*, vol. 30, 2017, Accessed: Jul. 27, 2025. [Online]. Available: <https://github.com/slundberg/shap>
- [12] M. Imani, A. Beikmohammadi, and H. R. Arabnia, “Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels,” *Technologies* 2025, Vol. 13, Page 88, vol. 13, no. 3, p. 88, Feb. 2025, doi: 10.3390/TECHNOLOGIES13030088.
- [13] E. Rohaeti and A. Andriyati, “Comparative Study of Predictive Classification Models on Data with Severely Imbalanced Predictors,” 2024.

-
- [14] E. C. Gök and M. O. Olgun, “SMOTE-NC and gradient boosting imputation based random forest classifier for predicting severity level of covid-19 patients with blood samples,” *Neural Comput Appl*, vol. 33, no. 22, pp. 15693–15707, Nov. 2021, doi: 10.1007/s00521-021-06189-y.
- [15] J. Friedman, T. Hastie, and R. Tibshirani, “Additive logistic regression: A statistical view of boosting,” 2000. doi: 10.1214/aos/1016218223.
- [16] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.
- [17] R. Diallo, C. Edalo, and O. O. Awe, “Machine Learning Evaluation of Imbalanced Health Data: A Comparative Analysis of Balanced Accuracy, MCC, and F1 Score,” *STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics and Health*, vol. Part F4005, pp. 283–312, 2025, doi: 10.1007/978-3-031-72215-8_12.
- [18] H. He and Y. Ma, *Imbalanced learning: Foundations, algorithms, and applications*, vol. 1. John Wiley & Sons, 2013. doi: 10.1002/9781118646106.
- [19] D. J. Hand, “A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems,” 2001.
- [20] M. Kuhn, D. Vaughan, and E. Hvitfeldt, “yardstick: Tidy Characterizations of Model Performance,” 2025. [Online]. Available: <https://github.com/tidymodels/yardstick>
- [21] N. Reviani and Y. E. Riany, “Establishing Healthy Eating Habits during Child Development to Reduce the Prevalence of Obesity,” 2022.
- [22] J. Jeon, S. Lee, and C. Oh, “Age-specific risk factors for the prediction of obesity using a machine learning approach.” [Online]. Available: <https://knhanes.kdca.go.kr/knhanes>
- [23] G. R. Patil and G. Sharma, “Overweight/obesity relationship with travel patterns, socioeconomic characteristics, and built environment,” *J Transp Health*, vol. 22, p. 101240, Sep. 2021, doi: 10.1016/J.JTH.2021.101240.
- [24] S. Yang *et al.*, “Access to fruit and vegetable markets and childhood obesity: A systematic review,” *Obesity Reviews*, vol. 22, no. S1, p. e12980, Feb. 2021, doi: 10.1111/OBR.12980.
- [25] F. Yuan *et al.*, “Association of Physical Activity and Sitting Time with Overweight/Obesity in Chinese Occupational Populations,” *Research Article Obes Facts*, vol. 14, pp. 141–147, 2021, doi: 10.1159/000512834.