

**RUMUSAN MASALAH
PROPOSAL RISET INFORMATIKA**

**“Penerapan Metode XGBoost dengan Optimasi SMOTE untuk Klasifikasi Tingkat
Obesitas Berdasarkan Gaya Hidup”**



Disusun Oleh:

Chesa Saskia Rafika / 22081010211

**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”
JAWA TIMUR
2025**

1. Latar Belakang

Obesitas merupakan salah satu permasalahan kesehatan global yang terus meningkat dari tahun ke tahun. Berdasarkan data World Health Organization (WHO), prevalensi obesitas telah meningkat lebih dari tiga kali lipat sejak tahun 1975 dan menjadi salah satu penyebab utama penyakit tidak menular seperti diabetes tipe 2, hipertensi, serta penyakit jantung (WHO, 2024). Di Indonesia, hasil Riset Kesehatan Dasar (Riskesdas) menunjukkan peningkatan angka obesitas baik pada laki-laki maupun perempuan dewasa setiap tahunnya (Kemenkes RI, 2023). Kondisi ini menunjukkan pentingnya upaya deteksi dini terhadap risiko obesitas untuk mencegah komplikasi lebih lanjut. Gaya hidup memiliki peran besar terhadap tingkat obesitas, termasuk pola makan, aktivitas fisik, durasi tidur, dan kebiasaan merokok. Identifikasi hubungan antara gaya hidup dengan tingkat obesitas dapat membantu dalam pengambilan keputusan medis dan edukasi masyarakat. Namun, proses identifikasi secara manual sulit dilakukan ketika jumlah data sangat besar dan bersifat kompleks.

Perkembangan teknologi *machine learning* memberikan solusi efektif untuk menganalisis pola dan memprediksi tingkat obesitas berdasarkan data gaya hidup. Salah satu algoritma yang banyak digunakan karena performanya yang tinggi adalah Extreme Gradient Boosting (XGBoost). Algoritma ini mampu menangani data dengan dimensi besar dan kompleksitas tinggi serta memiliki kemampuan generalisasi yang baik.

Namun, permasalahan utama dalam klasifikasi obesitas adalah ketidakseimbangan kelas (*imbalanced data*) misalnya, jumlah individu dengan obesitas berat jauh lebih sedikit dibandingkan kategori normal. Ketidakseimbangan ini dapat menurunkan akurasi model karena algoritma cenderung bias terhadap kelas mayoritas. Untuk mengatasi hal ini, digunakan metode SMOTE (Synthetic Minority Oversampling Technique) yang bekerja dengan menambah sampel sintetis pada kelas minoritas sehingga distribusi data menjadi lebih seimbang. Dengan demikian, penelitian ini berfokus pada penerapan metode XGBoost dengan optimasi SMOTE untuk meningkatkan performa klasifikasi tingkat obesitas berdasarkan gaya hidup. Diharapkan, hasil penelitian ini dapat menjadi kontribusi dalam bidang kesehatan prediktif berbasis *machine learning* serta memberikan insight terhadap faktor gaya hidup yang paling berpengaruh terhadap obesitas.

2. Rumusan Masalah

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana penerapan metode XGBoost untuk melakukan klasifikasi tingkat obesitas berdasarkan faktor gaya hidup?
2. Bagaimana pengaruh penerapan SMOTE terhadap peningkatan performa model klasifikasi obesitas?

3. Seberapa baik akurasi, presisi, dan *recall* yang dihasilkan oleh model XGBoost dengan optimasi SMOTE dibandingkan dengan tanpa SMOTE?

3. Identifikasi Masalah

Dari uraian di atas, dapat diidentifikasi beberapa permasalahan utama, yaitu:

1. Belum diketahui seberapa besar pengaruh setiap faktor gaya hidup terhadap tingkat obesitas.
2. Diperlukan model machine learning yang mampu menangani data tidak seimbang dan memiliki performa tinggi dalam klasifikasi.
3. Ketidakseimbangan jumlah data antar kelas tingkat obesitas menyebabkan hasil klasifikasi kurang optimal.
4. Belum ada penelitian yang secara khusus menerapkan kombinasi XGBoost dan SMOTE untuk kasus klasifikasi obesitas berbasis gaya hidup di konteks data Indonesia.

4. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Menerapkan metode **XGBoost** untuk klasifikasi tingkat obesitas berdasarkan data gaya hidup individu.
2. Mengoptimalkan hasil klasifikasi dengan menggunakan teknik **SMOTE** untuk menyeimbangkan distribusi data.
3. Menganalisis performa model dengan membandingkan hasil sebelum dan sesudah penerapan SMOTE.
4. Mengetahui faktor gaya hidup yang paling berpengaruh terhadap tingkat obesitas berdasarkan hasil interpretasi model XGBoost.