

BAB I

PENDAHULUAN

1.1 Latar Belakang

Obesitas adalah masalah kesehatan global yang terus mengalami peningkatan. Menurut Organisasi Kesehatan Dunia (WHO), pada tahun 2022 lebih dari 2,5 miliar orang dewasa di seluruh dunia mengalami kelebihan berat badan, dengan sekitar 890 juta orang di antaranya hidup dengan obesitas[1]. Kondisi ini telah meningkatkan risiko berbagai penyakit kronis, seperti diabetes tipe 2, penyakit jantung, dan hipertensi[2]. Yang mana hal ini menjadi beban serius bagi sistem kesehatan masyarakat di banyak negara.

Perubahan gaya hidup modern yang meliputi pola makan yang tidak sehat dan kurangnya aktivitas fisik menjadi penyebab utama. Dengan kebiasaan konsumsi makanan tinggi kalori, rendah gizi, serta meningkatnya konsumsi makanan cepat saji, berkontribusi besar pada tingginya prevalensi obesitas di berbagai kalangan mulai dari anak-anak, remaja, hingga orang dewasa[3].

Untuk mengatasi masalah obesitas, diperlukan langkah pencegahan yang didukung oleh alat klasifikasi status obesitas yang akurat. Model pembelajaran mesin (machine learning) telah terbukti efektif sebagai model klasifikasi status obesitas, salah satunya adalah algoritma Random Forest[4]. Namun, meskipun metode ini mampu memberikan hasil yang baik, diperlukan optimasi lebih lanjut untuk meningkatkan akurasi klasifikasi status obesitas pada individu.

Optimasi dapat dilakukan melalui pengelompokan variabel dengan sebaran luas (binning), Feature scaling dan Data Augmentation, Hyperparameter tuning menggunakan GridSearchCV, penerapan voting ensemble model dengan kombinasi Random Forest dan XGBoost, serta optimasi lebih lanjut melalui hyperparameter tuning pada model ensemble tersebut dapat menjadi langkah yang patut dicoba dalam upaya meningkatkan keakuratan model klasifikasi status obesitas. Sehingga hasilnya dapat memberikan kontribusi nyata dalam pencegahan dan penanganan obesitas secara lebih efektif di masa depan.

1.2 Rumusan Masalah

Apakah upaya penerapan metode Binning pada variabel, Feature Scaling, Data Augmentation, Hyperparameter Tuning, serta penerapan model ensemble antara Random Forest dan XGBoost dapat membantu meningkatkan akurasi model klasifikasi status obesitas?

1.3 Batasan Masalah

Penelitian ini difokuskan pada optimasi akurasi model klasifikasi status obesitas dengan menerapkan metode Binning pada variabel dengan sebaran nilai yang luas, Feature Scaling, Data Augmentation, Hyperparameter Tuning, penerapan voting ensemble model dengan kombinasi Random Forest dan XGBoost, serta optimasi lebih lanjut melalui hyperparameter tuning pada model ensemble tersebut. Adapun batasan penelitian ini meliputi:

1. Fokus pada Akurasi Model

Penelitian ini hanya membahas aspek optimasi akurasi model klasifikasi obesitas. Fokus akan diberikan pada metode Binning, Feature Scaling, Data Augmentation, serta Hyperparameter Tuning pada algoritma Random Forest dan XGBoost. Aspek lain seperti interpretabilitas model, efisiensi komputasi, dan implementasi model di dunia nyata tidak menjadi fokus utama.

2. Variabel yang Digunakan

Penelitian akan menggunakan seluruh variabel dalam dataset, dengan rincian sebagai berikut:

- a) Variabel yang digunakan merupakan seluruh variabel yang terdapat dalam dataset yang digunakan. Adapun rincian variabel yang digunakan meliputi
- b) Age (Umur dalam bentuk numerik integer)
- c) Height (Umur dalam bentuk numerik integer dengan skala cm)
- d) Overweight_Obese_Family (Riwayat keluarga Obesitas, data berisi boolean ya atau tidak)
- e) Consumption_of_Fast_Food (Konsumsi makanan cepat saji, data berisi boolean ya atau tidak)

- f) *Frequency_of_Consuming_Vegetables* (Frekuensi konsumsi sayuran, data berisikan tingkatan kategori Rarely, Sometimes, Always)
 - g) *Number_of_Main_Meals_Daily* (Jumlah makan utama dalam sehari)
 - h) *Food_Intake_Between_Meals* (Makanan yang dikonsumsi antara makan utama)
 - i) *Smoking* (Merokok atau tidak)
 - j) *Liquid_Intake_Daily* (Konsumsi Cairan harian)
 - k) *Calculation_of_Calorie_Intake* (Menghitung Masukan Kalori)
 - l) *Physical_Exercise* (Aktivitas/Latihan Fisik)
 - m) *Schedule_Dedicated_to_Technology* (Jadwal Menggunakan Teknologi)
 - n) *Type_of_Transportation_Used* (Transportasi yang digunakan)
 - o) *Class* (Klasifikasi apakah tergolong Underweight, Normal, Overweight, atau Obesitas)
3. **Faktor Lain di Luar Dataset**
Faktor eksternal yang tidak tercantum dalam dataset, tetapi berpotensi mempengaruhi obesitas, tidak akan dibahas dalam penelitian ini.
4. **Teknik Optimasi Model**
Teknik optimasi terbatas pada penggunaan Feature Scaling, Binning, Data Augmentation, Hyperparameter Tuning dengan GridSearchCV, serta penerapan voting ensemble model antara Random Forest dan XGBoost.
5. **Populasi Data**
Penelitian ini terbatas pada data dalam dataset yang tersedia. Generalisasi hasil ke populasi lebih luas membutuhkan pengujian tambahan dengan dataset yang lebih beragam.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah meningkatkan akurasi klasifikasi status obesitas dengan menerapkan metode Binning pada variabel dengan sebaran nilai

yang luas, Feature Scaling, Data Augmentation, serta Hyperparameter Tuning. Penelitian ini juga bertujuan untuk membuktikan bahwa kombinasi algoritma Random Forest dan XGBoost melalui voting ensemble model dapat menghasilkan akurasi yang lebih tinggi dibandingkan penelitian sebelumnya, yang mencapai 87,82%.

1.5 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi nyata dalam upaya pencegahan dan penanganan obesitas secara lebih efektif di masa depan. Dengan model klasifikasi yang lebih akurat, hasil penelitian ini dapat membantu para ahli kesehatan dalam mengambil keputusan berbasis data untuk intervensi yang lebih tepat sasaran.

1.6 Sistematika Penulisan

Berikut ini adalah sistematika penulisan skripsi pada penelitian ini:

BAB I PENDAHULUAN

Berisi Latar belakang masalah, rumusan masalah, Batasan masalah, tujuan penelitian manfaat penelitian dan sistematika penulisan skripsi.

BAB II TINJAUAN PUSTAKA

Berisi tinjauan pustaka, dan dasar-dasar teori yang digunakan.

BAB III METODE PENELITIAN

Didalamnya terdapat tinjauan umum tentang alur penelitian, analisis masalah, hingga solusi yang ditawarkan.

BAB IV HASIL DAN PEMBAHASAN

Di bab ini penulis akan menyimpulkan hasil penelitian serta pembahasannya.

BAB V PENUTUP

Berisi kesimpulan dan saran yang dapat peneliti rangkum selama proses penelitian.