

BAB I

PENDAHULUAN

1.1 Latar Belakang

Obesitas telah menjadi masalah kesehatan global yang semakin meningkat dan menjadi ancaman serius. Berdasarkan data WHO pada 2022, 1 dari 8 orang di dunia mengalami obesitas, dengan prevalensi obesitas pada orang dewasa meningkat dua kali lipat sejak 1990, sementara obesitas pada remaja meningkat empat kali lipat[1]. Obesitas dapat menyebabkan peningkatan resiko berbagai penyakit kronis, seperti diabetes tipe 2, penyakit jantung, dan kanker yang dapat menurunkan kualitas hidup individu[1]. Selain dampak kesehatan, obesitas juga memberikan dampak ekonomi yang besar dengan meningkatnya biaya perawatan medis dan penurunan produktivitas kerja akibat penyakit terkait obesitas[2][3].

Pola hidup sehari-hari memiliki pengaruh besar terhadap peningkatan obesitas. Kebiasaan makan yang buruk, seperti mengonsumsi makanan tinggi kalori dengan gizi rendah serta pola makan yang tidak teratur, meningkatkan resiko seseorang terkena obesitas. Kurangnya aktivitas fisik juga menjadi faktor penting yang berkontribusi pada obesitas, demikian pula dengan kurang tidur yang dapat mempengaruhi metabolisme tubuh dan meningkatkan nafsu makan[4][5][6]. Oleh karena itu, menjaga pola hidup sehat dengan makan makanan bergizi seimbang, rutin berolahraga, dan tidur yang cukup sangat penting dalam mencegah obesitas dan menjaga kesehatan tubuh secara keseluruhan. Deteksi dini terhadap obesitas sangat krusial agar penderita dapat segera mendapatkan penanganan yang sesuai untuk mencegah komplikasi lebih lanjut dan meningkatkan kualitas hidup mereka.

Seiring dengan berkembangnya teknologi, algoritma *machine learning* dapat digunakan untuk memprediksi tingkat obesitas berdasarkan pola hidup sehari-hari, yang menjadi upaya penting dalam pencegahan dini. Extreme Gradient Boosting (XGBoost) telah terbukti efektif dengan memberikan akurasi tinggi, mencapai 97,16%[7] dan 99,61%[8] dalam beberapa penelitian. Meskipun memiliki keunggulan dalam menangani dataset besar dan kompleks, XGBoost

memerlukan waktu pelatihan yang lebih lama dan sumber daya komputasi yang lebih besar. Di sisi lain, Logistic Regression juga merupakan algoritma yang efisien dan mudah dipahami, dengan akurasi yang mencapai 99,05%[9], 99,27%[10], dan 98,99%[11] dalam penelitian sebelumnya. Namun, Logistic Regression memiliki keterbatasan dalam menangani data yang sangat kompleks atau non-linear.

Penelitian ini menggunakan dataset "Obesity Prediction Dataset" untuk menganalisis faktor-faktor gaya hidup yang mempengaruhi obesitas, seperti kebiasaan makan, aktivitas fisik, kebiasaan merokok, konsumsi alkohol, penggunaan teknologi, serta jenis transportasi. Penelitian ini juga mempertimbangkan faktor demografis seperti jenis kelamin, umur, tinggi badan, berat badan, riwayat obesitas keluarga, dan status pemantauan kalori. Dengan menggunakan teknologi *machine learning*, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan sistem kesehatan yang lebih efektif dalam mendiagnosis dan mengobati obesitas pada tahap awal. Variabel obesitas akan digunakan untuk mengidentifikasi tingkat obesitas responden, mulai dari kekurangan berat badan hingga berbagai tingkat obesitas.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, rumusan masalah dalam penelitian ini yaitu:

1. Bagaimana kinerja Extreme Gradient Boost dan Logistic Regression dalam memprediksi tingkat obesitas berdasarkan pola hidup sehari-hari?
2. Apa saja kelebihan dan kekurangan dari masing-masing algoritma dalam konteks prediksi obesitas?

1.3 Batasan Masalah

Batasan masalah dari penelitian ini yaitu:

1. Penelitian ini menggunakan dataset "Obesity Prediction Dataset" yang didapat dari Kaggle.

2. Penelitian ini menganalisis dua algoritma *machine learning*, yaitu Extreme Gradient Boost dan Logistic Regression, tanpa membandingkan algoritma lain.
3. Penelitian ini akan menganalisis faktor-faktor gaya hidup dan kebiasaan yang mempengaruhi status obesitas, dengan fokus pada variabel seperti kebiasaan makan (makanan tinggi kalori, frekuensi konsumsi sayuran, camilan, dan konsumsi air), aktivitas fisik, kebiasaan merokok dan konsumsi alkohol, penggunaan teknologi, serta jenis transportasi. Selain itu, penelitian juga akan mempertimbangkan faktor demografis (jenis kelamin, umur, tinggi dan berat badan), riwayat obesitas dalam keluarga, serta status pemantauan kalori. Variabel Obesity akan digunakan sebagai target untuk mengidentifikasi tingkat obesitas responden, mulai dari kekurangan berat badan hingga berbagai tingkat obesitas.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai oleh peneliti dalam penelitiannya adalah

1. Untuk membandingkan kinerja Extreme Gradient Boost dan Logistic Regression dalam memprediksi obesitas berdasarkan data pola hidup sehari-hari.
2. Untuk menilai akurasi, kelebihan dan kekurangan masing-masing algoritma dalam memprediksi status obesitas, serta menemukan algoritma paling efektif digunakan dalam prediksi obesitas.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah memberikan pemahaman mengenai keunggulan dan keterbatasan algoritma Extreme Gradient Boosting (XGBoost) dan Logistic Regression dalam memprediksi tingkat obesitas berdasarkan pola hidup sehari-hari, sehingga dapat memperkaya literatur dalam bidang *machine learning* dalam prediksi kesehatan.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Pendahuluan berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian dan manfaat penelitian.

BAB II TINJAUAN PUSTAKA

Tinjauan pustaka berisi penelitian sebelumnya yang menjadi acuan dan teori-teori yang digunakan dalam penelitian ini.

BAB III METODE PENELITIAN

Metode penelitian menguraikan menguraikan desain penelitian, data dan sumber data yang digunakan, teknik pengumpulan data, serta metode analisis yang diterapkan untuk membandingkan kinerja XGBoost dan Logistic Regression.

BAB IV HASIL DAN PEMBAHASAN

Hasil dan pembahasan menyajikan hasil yang diperoleh dari penerapan kedua model, dilanjutkan dengan pembahasan mengenai hasil perbandingan kinerja antara XGBoost dan Logistic Regression, serta analisis faktor-faktor yang mempengaruhi performa.

BAB V PENUTUP

Penutup berisi kesimpulan yang diambil berdasarkan hasil penelitian dan saran untuk penelitian selanjutnya .