

ABSTRAK

Obesitas merupakan salah satu permasalahan kesehatan yang semakin meningkat dan dapat dipengaruhi oleh berbagai faktor seperti pola makan, aktivitas fisik, serta kondisi antropometri seseorang. Oleh karena itu, diperlukan metode yang mampu melakukan analisis dan prediksi tingkat obesitas secara efektif. Penelitian ini bertujuan untuk memprediksi tingkat obesitas berdasarkan faktor gaya hidup dan antropometri menggunakan metode machine learning yaitu Logistic Regression dan XGBoost. Dataset yang digunakan dalam penelitian ini diperoleh dari platform Kaggle yang terdiri dari 2111 data dengan 17 atribut yang mencakup karakteristik individu, kebiasaan gaya hidup, serta indikator kondisi fisik. Proses penelitian diawali dengan tahap preprocessing data yang meliputi pemeriksaan struktur data, pengecekan missing value, serta transformasi data kategorikal menggunakan Label Encoding. Dataset kemudian dipisahkan menjadi fitur dan label dengan variabel NObeyesdad sebagai target klasifikasi. Selanjutnya dilakukan pembagian dataset menggunakan metode train-test split dengan rasio 80% data latih dan 20% data uji, serta normalisasi data menggunakan StandardScaler. Model kemudian dilatih menggunakan algoritma Logistic Regression dan XGBoost, kemudian dievaluasi menggunakan metrik accuracy, precision, recall, dan F1-score. Hasil penelitian menunjukkan bahwa model Logistic Regression memperoleh nilai accuracy sebesar 87,47%, sedangkan model XGBoost memperoleh nilai accuracy sebesar 95,98%. Berdasarkan hasil perbandingan tersebut, model XGBoost menunjukkan performa yang lebih baik dalam memprediksi tingkat obesitas dibandingkan Logistic Regression.

Kata kunci: Obesitas, Machine Learning, Logistic Regression, XGBoost, Prediksi.