

ABSTRAK

Diabetes melitus (DM) merupakan penyakit kronis yang terjadi ketika tubuh tidak mampu memproduksi insulin secara optimal atau tidak dapat menggunakan insulin secara efektif. Diabetes dapat mempengaruhi berbagai sistem organ tubuh manusia dalam jangka waktu tertentu yang disebut komplikasi. Berdasarkan data dari Puskesmas Padangsari, terdapat minimal 4 penderita diabetes, hipertensi, maupun keduanya di setiap rukun tetangga wilayah Padangsari dan Pedalangan.

Penelitian ini bertujuan mengembangkan model machine learning berbasis data non-invasif untuk klasifikasi dini risiko diabetes sebagai sistem pendukung keputusan bagi tenaga kesehatan. Metode yang digunakan adalah ensemble learning dengan algoritma AdaBoost dan Gradient Boosting. Prapemrosesan dilakukan sebelum pelatihan, dimulai dari balancing menggunakan SMOTE, mapping, analisis outlier, encoding, dan pemisahan data latih dan uji. Pelatihan model AdaBoost menggunakan algoritma SAMME.R yang memanfaatkan estimasi probabilitas kelas dari weak learner untuk memperbarui bobot, sedangkan Gradient Boosting membuat objek model dengan parameter default dan melakukan proses training. Kedua model kemudian memprediksi kelas pada data uji dan hasilnya disimpan.

Evaluasi dilakukan menggunakan Classification Report, Confusion Matrix, ROC, PR Curve, dan Feature Importance. Hasil menunjukkan Gradient Boosting dan AdaBoost memiliki performa hampir identik dengan akurasi 94%. Pada kelas Tidak Diabetes, keduanya sangat baik karena memiliki recall mendekati sempurna. Namun pada kelas Diabetes, performa masih kurang optimal dengan recall di bawah 0.5, menandakan sebagian kasus tidak terdeteksi. Secara agregat, AdaBoost sedikit lebih unggul pada macro average recall dan macro average F1-score, menandakan kemampuannya lebih baik dalam menyeimbangkan prediksi antar kelas.

Kata kunci: *Machine Learning; Diabetes Melitus; Klasifikasi; AdaBoost; Gradient Boost*