

ABSTRAK

Hipertensi merupakan salah satu penyakit tidak menular dengan prevalensi tinggi yang menjadi penyebab utama kematian global, sehingga deteksi dini dan klasifikasi status hipertensi menjadi hal yang penting. Penelitian ini bertujuan untuk membandingkan kinerja dua metode klasifikasi, yaitu Regresi Logistik Biner dan algoritma Naïve Bayes, dalam mengklasifikasikan status hipertensi pada pasien di Puskesmas Padangsari Kota Semarang. Kedua metode merupakan pendekatan *supervised learning* yang populer dalam data mining adalah Regresi Logistik Biner yang digunakan untuk mengidentifikasi pengaruh variabel independen terhadap status hipertensi, dan Naïve Bayes menghitung probabilitas klasifikasi berdasarkan asumsi independensi antar fitur. Penelitian dilakukan terhadap 600 data pasien tahun 2024 dengan berbagai proporsi data *training* dan testing, serta evaluasi menggunakan confusion matrix, sensitivitas, dan nilai AUC. Hasil penelitian menunjukkan bahwa pada proporsi 70%:30%, Regresi Logistik Biner memiliki kinerja terbaik dibandingkan proporsi lainnya, dengan akurasi pada data testing sebesar 67,39% dan AUC sebesar 72,4%, serta akurasi pada data *training* sebesar 67,57% dan AUC sebesar 70,8%. Selisih sensitivitas yang kecil antara data *training* dan *testing* (0,18%) menunjukkan kestabilan model dan ketidaktergantungan terhadap overfitting. Oleh karena itu, metode Regresi Logistik Biner dinilai lebih optimal dalam mengklasifikasikan status hipertensi pasien dibandingkan Algoritma Naive Bayes.

Kata Kunci: Hipertensi, Regresi Logistik Biner, Naïve Bayes, Klasifikasi, Akurasi, AUC.