

BAB V

PENUTUP

3.4 Kesimpulan

Pada penelitian ini terdapat 8 variabel input yaitu jenis kelamin, usia, hipertensi, penyakit jantung, riwayat merokok, indeks massa tubuh, kadar glukosa, kadar gula darah dan 1 variabel kelas. Dari 100.000 data setelah dilakukan *preprocessing* data menjadi 96.146 data.

Dari hasil proses Klasifikasi pemodelan algoritma *naive bayes* terhadap *Diabetes Prediction Dataset* menggunakan 3 perbandingan yaitu 90:10, 80:20 dan 50:50. Hasil akurasi tertinggi didapatkan pada perbandingan 80:20 dengan akurasi = 90,38%, presisi = 45,84%, recall = 65,73% dan f1 score = 54,01%.

3.5 Saran

Saran dari Penerapan Metode Naive Bayes Untuk Klasifikasi Penyakit Diabetes adalah sebagai berikut :

1. Terapkan algoritma yang dirancang untuk menangani data yang tidak seimbang, seperti *SMOTE (Synthetic Minority Over-Sampling Technique)* atau *cost-sensitive learning*.
2. Gunakan teknik validasi silang yang sesuai, seperti *stratified k-fold cross-validation*, untuk memastikan bahwa semua kelas terwakili secara proporsional dalam setiap lipatan.
3. Coba berbagai jenis metode dan parameter untuk menemukan kombinasi yang paling baik dalam menangani data tidak seimbang.