

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, penerapan algoritma *Naive Bayes* dan *Decision Tree* pada Pima Indians Diabetes Dataset dapat digunakan untuk melakukan prediksi status diabetes berdasarkan karakteristik data pasien. Penelitian ini melalui beberapa tahapan, yaitu preprocessing data, penanganan missing value, standardisasi fitur, serta pembagian data menggunakan metode stratified dengan rasio 80:20. Melalui penerapan metode supervised learning, tahapan tersebut mampu menghasilkan model klasifikasi yang dapat membedakan pasien dengan kondisi diabetes dan non-diabetes. Hasil pengujian menggunakan confusion matrix menunjukkan bahwa kedua algoritma dapat diterapkan sebagai metode prediksi penyakit diabetes berdasarkan fitur Age, Glucose, BloodPressure, Insulin, dan BMI.

Hasil analisis perbandingan performa algoritma *Naive Bayes* dan *Decision Tree* menunjukkan adanya perbedaan kemampuan klasifikasi berdasarkan nilai akurasi, sensitivitas, dan spesifisitas, baik pada pengujian baseline maupun setelah optimasi menggunakan Stratified K-Fold Cross Validation dan GridSearchCV. Penelitian ini menempatkan sensitivitas sebagai metrik utama untuk mengurangi risiko false negative pada kasus medis. Berdasarkan fokus tersebut, algoritma *Decision Tree* memiliki performa yang lebih baik dalam mendeteksi pasien yang berisiko mengalami diabetes, sehingga dipilih sebagai model yang digunakan dalam sistem karena kemampuan deteksinya lebih sesuai dengan kebutuhan skrining awal.

Implementasi model *machine learning* ke dalam sistem prediksi penyakit diabetes berbasis web berhasil dilakukan dengan menyediakan fitur penginputan data pasien, proses prediksi status diabetes, pengelolaan data pasien, serta penyajian hasil perbandingan performa algoritma. Sistem yang dikembangkan memiliki alur

penggunaan yang sederhana sehingga dapat membantu admin maupun tenaga kesehatan dalam melakukan skrining awal terhadap risiko diabetes. Hasil pengujian sistem menggunakan metode BlackBox Testing menunjukkan tingkat keberhasilan sebesar 100%, di mana seluruh fungsi utama, mulai dari login, penginputan data pasien, proses prediksi, penyimpanan data hasil prediksi, hingga komparasi algoritma, telah berjalan sesuai dengan rancangan yang diharapkan. Namun, sistem ini tetap berperan sebagai alat bantu pendukung keputusan dan tidak menggantikan proses diagnosis maupun keputusan medis yang dilakukan oleh tenaga kesehatan.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa hal yang dapat dikembangkan pada penelitian selanjutnya.

1. Penelitian berikutnya disarankan menggunakan data langsung dari pasien atau bekerja sama dengan instansi kesehatan agar data lebih sesuai kondisi nyata, serta menambah jumlah data dan variabel klinis lain seperti riwayat keluarga, kadar HbA1c, dan gaya hidup untuk meningkatkan kemampuan prediksi model.
2. Penelitian selanjutnya dapat memperluas kajian dengan membandingkan algoritma lain seperti *Random Forest*, *Support Vector Machine*, atau *XGBoost*, serta mengombinasikannya dengan teknik penanganan ketidakseimbangan kelas seperti *SMOTE* atau *class weighting* untuk menemukan metode terbaik dalam menghasilkan sensitivitas dan performa klasifikasi optimal.
3. Sistem prediksi berbasis web ini masih dapat disempurnakan melalui penambahan fitur seperti riwayat hasil prediksi, pengaturan hak akses pengguna, layanan prediksi mandiri bagi pasien, serta integrasi dengan rekam medis elektronik, untuk meningkatkan kemudahan penggunaan dan efektivitas sistem sebagai media skrining risiko diabetes