BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang berjudul "Implementasi Algoritma *Random Forest* untuk Prediksi Risiko Diabetes Berbasis Data Klinis", serta hasil analisis dan pembahasan yang telah dilakukan, maka dapat disimpulkan beberapa hal sebagai berikut:

1. Performa model *Random Forest* dalam membangun sistem prediksi terbukti akurat dan dapat diandalkan. Melalui proses *hyperparameter tuning* menggunakan *RandomizedSearchCV*, model memperoleh parameter terbaik dengan skor *F1-Score* rata-rata sebesar 0.967 pada *cross-validation* (5-*Fold*). Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dan mampu mempertahankan performa tinggi meskipun diuji pada data yang berbeda.
2. Faktor-faktor klinis yang paling berpengaruh terhadap prediksi risiko diabetes berdasarkan hasil analisis pada model *Random Forest* secara berturut-turut adalah Glukosa Darah Sewaktu, Glukosa Darah Puasa, Usia, Tekanan Darah Sistolik (*BP_SYSTOLIC*), dan Persentase kadar HbA1c. Faktor-faktor tersebut memiliki kontribusi tertinggi terhadap hasil prediksi, yang menandakan bahwa parameter kadar glukosa dan tekanan darah memiliki pengaruh dominan dalam menentukan risiko diabetes pada pasien.
3. Efektivitas algoritma *Random Forest* dalam mendeteksi risiko diabetes terbukti sangat baik. Berdasarkan evaluasi menggunakan 5-*Fold* Cross Validation, model menghasilkan rata-rata akurasi sebesar 93,6%, presisi 94,2%, *recall* 99,2%, dan *F1-Score* 0,967. Nilai *recall* yang mendekati sempurna menunjukkan bahwa model mampu mendeteksi hampir seluruh pasien DM dengan benar, yang sangat krusial dalam konteks medis karena

kegagalan mendeteksi pasien DM dapat berdampak serius terhadap penanganan pasien. Nilai ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data baru dan dapat diandalkan sebagai alat bantu deteksi dini.

4. Implementasi model ke dalam aplikasi berbasis web menggunakan *Streamlit* berjalan dengan sukses dan menghasilkan sistem yang interaktif serta mudah digunakan. Pengguna dapat memasukkan data klinis pasien seperti usia, tekanan darah, kadar glukosa, dan HbA1c untuk memperoleh hasil prediksi secara langsung. Berdasarkan *black box testing*, seluruh komponen aplikasi, mulai dari *input* data, pemanggilan model, perhitungan probabilitas, hingga tampilan hasil, berfungsi dengan baik dan sesuai dengan rancangan. Aplikasi ini dapat dijadikan alat bantu deteksi dini risiko diabetes yang praktis dan efisien.

5.2 Saran

Berdasarkan hasil penelitian yang telah diperoleh, maka penulis memberikan beberapa saran untuk pengembangan penelitian dan implementasi di masa yang akan datang, yaitu sebagai berikut:

1. Perluasan dan Pembaruan *Dataset* Disarankan untuk menggunakan *dataset* dengan jumlah sampel yang lebih besar dan lebih beragam, mencakup data dari berbagai wilayah, usia, serta faktor gaya hidup seperti aktivitas fisik dan riwayat keluarga, agar model dapat lebih adaptif terhadap variasi populasi.
2. Perbandingan dengan Algoritma Lain Penelitian selanjutnya dapat melakukan perbandingan dengan algoritma lain seperti *XGBoost*, *Gradient Boosting*, *Support Vector Machine* (SVM), atau *Neural Network* untuk menilai algoritma mana yang memberikan performa terbaik dalam mendeteksi risiko diabetes.
3. Peningkatan Transparansi Model (*Explainability*) Penggunaan metode *Explainable* AI (XAI) seperti SHAP atau LIME disarankan agar model dapat memberikan penjelasan lebih mendalam mengenai alasan di balik

setiap hasil prediksi, sehingga dapat meningkatkan kepercayaan pengguna dan tenaga medis terhadap sistem.

4. Aplikasi berbasis web yang telah dibangun dapat dikembangkan dengan menambahkan fitur seperti penyimpanan riwayat pasien, sistem autentikasi pengguna, *dashboard* analitik, dan integrasi ke dalam sistem informasi rumah sakit agar dapat digunakan dalam praktik klinis secara nyata.