

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai penerapan algoritma *Random Forest* untuk deteksi penyakit jantung, serta optimisasi melalui seleksi fitur menggunakan *feature importance*, maka dapat disimpulkan hal-hal berikut:

1. Algoritma *Random Forest* mampu membangun model klasifikasi penyakit jantung dengan akurasi yang tinggi, yaitu mencapai **82,61%** pada model awal yang menggunakan seluruh fitur dari dataset UCI Heart Disease.
2. Analisis *feature importance* berhasil mengidentifikasi fitur-fitur yang paling berpengaruh terhadap prediksi, di antaranya: *thalach*, *chol*, *age*, dan *oldpeak*.
3. Penghapusan fitur-fitur dengan kontribusi rendah seperti *id*, *dataset*, *ca*, dan *fbs* tidak hanya menyederhanakan model, tetapi juga meningkatkan efisiensi waktu pelatihan sebesar **15,6%** dan waktu prediksi sebesar **12,1%**.
4. Model dengan fitur yang diseleksi menunjukkan akurasi yang lebih tinggi, yaitu **84,78%**, dengan penurunan *ROC AUC* yang relatif kecil dari **0,9116** menjadi **0,9008**.
5. Dengan performa yang kompetitif dan efisiensi yang lebih baik, model *reduced features* dinilai lebih praktis untuk diimplementasikan sebagai sistem pendukung keputusan dalam deteksi dini penyakit jantung.

5.2 Saran

Untuk pengembangan penelitian selanjutnya, penulis menyarankan beberapa hal berikut:

1. Melakukan eksperimen tambahan menggunakan algoritma lain seperti *XGBoost*, *LightGBM*, atau *Support Vector Machine* untuk perbandingan performa.

2. Menerapkan teknik balancing data seperti *SMOTE* atau *undersampling* apabila ditemukan distribusi kelas yang tidak seimbang.
3. Menggunakan data medis dari sumber lain atau data klinis nyata untuk menguji generalisasi model terhadap konteks dunia nyata.
4. Mengembangkan antarmuka sistem diagnosis berbasis web atau aplikasi mobile yang mengintegrasikan model klasifikasi sebagai bagian dari sistem pendukung keputusan.
5. Menambahkan analisis interpretabilitas lanjutan menggunakan pendekatan seperti *SHAP* atau *LIME* untuk meningkatkan transparansi model dalam konteks medis.

