

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi bakteri *E. coli* menggunakan algoritma *Decision Tree* dan *Random Forest* dengan data dari *UCI Machine Learning Repository* yang dibagi secara proporsional menjadi 80% data latih dan 20% data uji, melalui tahapan pra pemrosesan, pelatihan, dan pengujian, diperoleh hasil bahwa dataset tidak memiliki *missing value* maupun duplikat, telah dinormalisasi dengan *Min-Max Scaler* serta dilakukan *label encoding*, dan berdasarkan evaluasi menggunakan akurasi, *precision*, *recall*, dan *F1-score*, algoritma *Random Forest* menunjukkan performa terbaik dengan akurasi 88.24% dan *F1-score* 88.64% dibandingkan *Decision Tree* dengan akurasi 79.41% dan *F1-score* 80.66%, sehingga *Random Forest* dinilai lebih efektif, stabil, mampu mengatasi *overfitting*, serta konsisten dalam klasifikasi multikelas data bakteri *E. coli*.

5.2 Saran

Berdasarkan hasil dan temuan dalam penelitian ini, beberapa saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Penelitian di masa depan dapat mengembangkan metode klasifikasi dengan mengeksplorasi algoritma lain seperti *XGBoost*, *SVM*, atau pendekatan berbasis *deep learning* guna meningkatkan akurasi dan efisiensi klasifikasi.
2. Disarankan untuk menerapkan teknik penyeimbangan data seperti *SMOTE* pada dataset *E. coli*, karena distribusi kelas yang tidak merata dapat mempengaruhi performa klasifikasi, terutama pada kelas minoritas.
3. Sistem klasifikasi ini dapat dikembangkan lebih lanjut dalam bentuk aplikasi berbasis web atau desktop untuk diterapkan di laboratorium mikrobiologi atau bidang kesehatan sebagai alat bantu diagnosis awal.
4. Evaluasi model dapat diperluas dengan menambahkan visualisasi seperti *ROC Curve*, analisis sensitivitas model terhadap fitur, serta studi terhadap pengaruh parameter *tuning* pada performa algoritma.