

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai penanganan *missing value* dengan pendekatan *Center of Tendency* pada klasifikasi potabilitas air minum menggunakan algoritma SVM, *Naïve Bayes*, dan *Decision Tree*, maka dapat disimpulkan sebagai berikut:

- 1) Metode imputasi *missing value* menggunakan pendekatan *Center of Tendency* (Mean, Median, dan Modus) dapat digunakan untuk mengatasi *missing value* dalam dataset potabilitas air minum. Namun, metode ini tidak selalu memberikan dampak signifikan terhadap peningkatan performa model klasifikasi, terutama pada algoritma yang bergantung pada support vectors seperti SVM.
- 2) Hasil evaluasi menunjukkan bahwa penerapan metode imputasi *missing value* dengan pendekatan *Center of Tendency* memiliki pengaruh yang bervariasi terhadap akurasi model. Pada algoritma SVM, tidak terdapat peningkatan akurasi setelah dilakukan imputasi, dengan nilai akurasi tetap sebesar 61%. Pada algoritma *Naïve Bayes*, terjadi peningkatan akurasi dengan hasil terbaik sebesar 62% pada metode mean dan median. Sedangkan pada algoritma *Decision Tree*, metode median memberikan hasil terbaik dengan akurasi sebesar 58%.
- 3) Dari hasil evaluasi yang telah dilakukan, dapat disimpulkan bahwa metode imputasi median memberikan hasil terbaik dalam meningkatkan performa model *Decision Tree*, sedangkan pada *Naïve Bayes*, metode mean dan median menghasilkan peningkatan performa yang serupa. Sementara itu, pada algoritma SVM, metode imputasi tidak memberikan peningkatan yang signifikan terhadap performa model karena SVM lebih bergantung pada support vectors dibandingkan dengan distribusi data secara keseluruhan.

Dengan demikian, penggunaan pendekatan *Center of Tendency* dalam mengatasi *missing value* pada dataset potabilitas air minum dapat memberikan sedikit peningkatan performa pada beberapa algoritma klasifikasi tertentu, namun tidak selalu berdampak positif terhadap semua model yang digunakan.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan lebih lanjut, baik dari segi metode maupun implementasi penelitian:

- 1) Dalam Penelitian ini, model yang digunakan adalah SVM, Naive Bayes, dan Decision Tree. Untuk penelitian di masa depan dapat mencoba beberapa algoritma yang lain seperti Random Forest, XGBoots, atau Neural Networks, yang mungkin lebih adaptif terhadap metode imputasi Center of Tendency.
- 2) Penelitian ini hanya menggunakan Confusion Matrix sebagai teknik evaluasi. Untuk penelitian selanjutnya, disarankan untuk menggunakan teknik evaluasi k-fold cross-validation agar hasil evaluasi lebih stabil dan tidak bergantung pada satu pembagian data saja.