

BAB V

PENUTUP

5.1 KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut:

1. Penelitian ini menggunakan metode *Random Forest*, *Support Vector Machine (SVM)*, dan *K-Nearest Neighbors (KNN)* kemudian diimplementasikan dengan melakukan *pre-processing*, pemilihan fitur, pengujian model dan evaluasi hasil untuk mengidentifikasi kinerja model dalam mengklasifikasi penyakit ginjal kronis.
2. Penelitian ini menguji metode *Random Forest*, *Support Vector Machine (SVM)* dan *K-Nearest Neighbors (KNN)* dengan melakukan 10 kali pengujian dengan mengambil nilai tertinggi dari setiap metode. Selanjutnya, diterapkan pemilihan fitur *Recursive Feature Elimination (RFE)* dengan menguji 6, 12, dan 18 fitur, serta melakukan evaluasi *10-Fold Cross Validation* dengan mengambil nilai rata-rata *accuracy* agar didapat hasil yang lebih stabil. Didapatkan hasil *accuracy* tertinggi oleh metode *Random Forest* dan *Support Vector Machine (SVM)* tanpa menggunakan pemilihan fitur didapat *accuracy* 100%. Sementara itu, metode *K-Nearest Neighbors (KNN)* mencapai *accuracy* 98% dengan penggunaan 12 fitur. Selanjutnya, dengan penerapan *10-Fold Cross Validation* berdasarkan nilai rata-rata *accuracy*, metode *Random Forest* menunjukkan kinerja tertinggi dengan penggunaan 18 fitur, menghasilkan

accuracy sebesar 99.2%. Metode *Support Vector Machine (SVM)* memperoleh kinerja terbaik dengan menggunakan 12 fitur, mencapai *accuracy* sebesar 97.2%, sedangkan *K-Nearest Neighbors (KNN)* mencapai kinerja terbaik dengan menggunakan 6 fitur dan mendapatkan *accuracy* 95%. Berdasarkan hasil penelitian, metode yang paling baik dalam mengklasifikasi penyakit ginjal kronis adalah *Random Forest*.

5.2 SARAN

Penelitian ini disadari oleh penulis bahwa masih ada beberapa kekurangan yang menjadi saran atau masukan untuk menyempurnakan penelitian ini adalah sebagai berikut :

1. Penelitian dapat dikembangkan dengan menggunakan *dataset* yang lebih besar dan lebih kompleks. Penggunaan *dataset* yang lebih luas dapat memberikan pemahaman yang lebih mendalam tentang kemampuan model dalam menangani variasi data yang lebih kompleks.
2. Melakukan variasi pada pembagian data, seperti 80:20, 70:30, dan 60:40, pada penelitian selanjutnya. Hal ini dimaksud untuk memperoleh pemahaman yang lebih komprehensif tentang kinerja model dalam skenario yang berbeda.