

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian, pengujian, dan analisis yang telah dilakukan terhadap data Penyakit Ginjal Kronis (*Chronic Kidney Disease / CKD*), maka dapat ditarik kesimpulan sebagai berikut:

1. Algoritma *K-Nearest Neighbor* (KNN) terbukti mampu melakukan klasifikasi status CKD dengan performa yang baik, khususnya setelah dilakukan optimasi *hyperparameter*. Pada pengujian menggunakan parameter *default*, algoritma KNN menghasilkan nilai *accuracy* sebesar 97,5%, *precision* 100%, *recall* 96%, dan *F1-score* 97,96%. Setelah dilakukan *hyperparameter tuning* menggunakan metode *Grid Search*, performa KNN mengalami peningkatan yang signifikan, terutama pada metrik *recall* yang meningkat menjadi 98%, dengan *accuracy* 98,8% dan *F1-score* 99%. Peningkatan nilai *recall* ini menunjukkan bahwa model semakin mampu mendeteksi hampir seluruh pasien yang benar-benar menderita CKD, sehingga risiko kesalahan *false negative* dapat ditekan.
2. Algoritma *Naive Bayes* juga menunjukkan performa klasifikasi yang stabil dan cukup tinggi dalam mendeteksi CKD, baik sebelum maupun sesudah proses optimasi. Pada konfigurasi parameter *default*, algoritma *Naive Bayes* menghasilkan nilai *accuracy* 97,5%, *precision* 100%, *recall* 96%, dan *F1-score* 98%. Setelah dilakukan *hyperparameter optimization* menggunakan *Grid Search* dengan parameter *var_smoothing*, nilai evaluasi yang dihasilkan tetap berada pada tingkat yang sama, yaitu *accuracy* 97,5% dan *recall* 96%. Hasil ini menunjukkan bahwa optimasi *hyperparameter* tidak memberikan peningkatan signifikan pada algoritma *Naive Bayes*, yang mengindikasikan bahwa model tersebut sudah cukup optimal pada konfigurasi awal.
3. Berdasarkan perbandingan kinerja kedua algoritma dengan penekanan utama pada metrik *recall*, algoritma *K-Nearest Neighbor* dengan

hyperparameter tuning menggunakan *Grid Search* merupakan model terbaik dalam penelitian ini. Nilai *recall* tertinggi yang dicapai KNN sebesar 98% menunjukkan kemampuan model yang lebih unggul dalam mengidentifikasi pasien CKD dibandingkan *Naive Bayes* yang memiliki *recall* 96%. Dalam konteks kesehatan, tingginya nilai *recall* sangat penting karena berkaitan langsung dengan kemampuan model dalam mendeteksi kasus penyakit secara tepat dan meminimalkan kesalahan diagnosis berupa *false negative*, sehingga algoritma KNN hasil optimasi dinilai paling sesuai untuk mendukung proses deteksi dini Penyakit Ginjal Kronis.

5.2 Saran

Untuk pengembangan penelitian selanjutnya serta guna mengatasi keterbatasan yang masih terdapat dalam penelitian ini, disarankan agar dilakukan penambahan dan perluasan variasi fitur yang digunakan, khususnya fitur klinis dan laboratorium tambahan yang berkaitan dengan fungsi ginjal, riwayat penyakit penyerta, serta faktor gaya hidup pasien, sehingga model prediksi Penyakit Ginjal Kronis (*Chronic Kidney Disease/CKD*) dapat bekerja secara lebih komprehensif. Selain itu, penggunaan dataset dengan jumlah data yang lebih besar, lebih beragam, serta berasal dari berbagai institusi atau periode waktu yang berbeda perlu dipertimbangkan agar kemampuan generalisasi model semakin baik dan lebih representatif terhadap kondisi nyata. Penelitian selanjutnya juga disarankan untuk mengeksplorasi algoritma *machine learning* lain yang lebih kompleks, seperti metode berbasis *ensemble* dan *boosting* (misalnya *XGBoost*, *LightGBM*, atau *CatBoost*), serta melakukan perbandingan dengan pendekatan *deep learning* guna memperoleh model prediksi CKD yang memiliki performa dan stabilitas yang lebih optimal, khususnya dalam meningkatkan nilai *recall* sebagai metrik utama pada konteks kesehatan.