

BAB V

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan penelitian yang dilakukan dalam penerapan algoritma klasifikasi yaitu *Random Forest*, *C4.5*, *AdaBoost*, dan hibrida *C4.5-AdaBoost* untuk memprediksi kemungkinan seseorang mengalami penyakit hepatitis, dapat menarik beberapa kesimpulan penting. Penelitian ini bertujuan untuk membandingkan tingkat akurasi dari masing-masing algoritma tersebut, baik ketika digunakan tanpa optimalisasi maupun dengan optimalisasi, dalam memprediksi penyakit hepatitis

Berdasarkan hasil pengujian dan evaluasi yang dilakukan, dalam konteks prediksi penyakit hepatitis, recall menjadi fokus utama karena penting untuk memastikan deteksi yang tepat terhadap kasus positif. Ditemukan bahwa tingkat recall yang khusus memprediksi suatu penyakit positif dari enam pemodelan algoritma tanpa optimalisasi adalah sebagai berikut:

- Algoritma *Random Forest* sebesar 100%,
- Algoritma *C4.5* sebesar 96.88%,
- Algoritma *AdaBoost* sebesar 100%, dan
- Hibrida *C4.5-AdaBoost* sebesar 96.88%.

Sedangkan untuk tingkat recall dari enam pemodelan algoritma menggunakan optimalisasi sebagai berikut:

- Algoritma *Random Forest* sebesar 100%,
- Algoritma *C4.5* sebesar 100%,
- Algoritma *AdaBoost* sebesar 100%, dan
- Algoritma hibrida *C4.5-AdaBoost* sebesar 100%.

Hasil ini menunjukkan tingkat dari setiap pemodelan menunjukan perfoma yang sangat baik dan sempurna.

Namun berdasarkan ROC, presentase tingkat pemodelan dengan menggunakan optimalisasi pada setiap pemodelan menunjukan hasil yang sempurna sebesar 100% dan presentasi tingkat pada pemodelan *Random Forest* tanpa optimalisasi menunjukan hasil yang sempurna juga yaitu 100%

Berdasarkan hasil tersebut, dapat disimpulkan bahwa penggunaan algoritma pemodelan tanpa optimalisasi peneliti menyarankan menggunakan *Random Forest* berdasarkan tingkat ROC yang sempurna, dan keempat pemodelan dengan menggunakan optimalisasi dapat menjadi pilihan lainnya, dikarenakan setiap pemodelan memiliki tingkat yang sempurna.

5.2 Saran

Meskipun penelitian ini telah menghasilkan temuan yang menarik dan menjanjikan dalam penerapan algoritma klasifikasi hibrida *C4.5-AdaBoost* untuk memprediksi penyakit hepatitis, ada beberapa rekomendasi yang dapat diajukan untuk pengembangan lebih lanjut. Berikut adalah beberapa saran yang dapat dipertimbangkan untuk penelitian selanjutnya:

1. Penggunaan data yang terkini: Salah satu syarat penelitian ini adalah sebuah data yang akan digunakan yang memiliki batas waktu tertentu berdasarkan syarat penelitian dan tujuan yang ingin dicapai. Dengan dataset yang terkini, diharapkan dapat memberikan performa dan akurasi yang lebih baik.
2. Penggunaan dataset yang banyak: Penelitian ini menghadapi salah satu tantangan penting yaitu jumlah dataset yang kurang. Dikarenakan, dataset yang didapat peneliti tidak lengkap pada hasil cek lab yang tersimpan pada website instalasi rekam medis, sehingga membuat data yang didapatkan terdapat banyak *missing value* yang memungkinkan pemodelan masih kurang maksimal. Untuk penelitian selanjutnya, disarankan untuk dapat mengumpulkan hasil cek lab dan cek hati dengan minimal total 350 *row* data atau lebih dari satu rumah sakit, serta pengambilan data dari 1-2 bulan dari mulainya penelitian, dikarenakan banyaknya tahapan administrasi yang harus dilalui dan diselesaikan sebelum diizinkan pengambilan dataset.

Ketika dataset yang didapat lebih dari yang disarankan, hal ini akan memberikan nilai akurasi yang lebih tinggi dan lebih baik.

3. Penanganan *Missing Value* yang Lebih Efektif: Penelitian ini menghadapi salah satu tantangan lainnya dalam penanganan missing value pada data yang digunakan. Untuk penelitian selanjutnya, disarankan untuk mengimplementasikan strategi penanganan *missing value* yang lebih efektif. Hal ini akan membantu dalam meminimalkan pengaruh data yang tidak lengkap terhadap pemodelan dan meningkatkan akurasi prediksi.
4. Percobaan gabungan dengan Algoritma Lain: Meskipun penelitian ini secara khusus mengeksplorasi penerapan algoritma klasifikasi *Random Forest*, *C4.5*, *AdaBoost* dan hibrida *C4.5-AdaBoost*, penelitian yang akan datang dapat mempertimbangkan untuk melakukan eksperimen dengan menggabungkan algoritma lain. Penggabungan berbagai algoritma bisa memberikan manfaat tambahan dalam meningkatkan akurasi dan kehandalan prediksi.
5. Implementasi ke dalam sebuah Website: Meskipun pada penelitian hanya sampai pada pengetahuan untuk mengetahui pemodelan dan kinerja dari keempat model algoritma, diharapkan pada penelitian selanjutnya dapat mengimplementasikan kedalam program website dan dapat diuji coba dengan data yang sesungguhnya, dengan begitu dapat memberikan manfaat secara langsung dalam memprediksi pasien terkena penyakit hepatitis.