BAB V
PENUTUP

5.1 Kesimpulan

Dalam skripsi ini, penulis memilih metode *K-Nearest Nei*ghbors (KNN) dan *Random Forest* untuk tugas klasifikasi. Penulis memilih *Random Forest* karena kemampuannya mengatasi *overfitting* dan variabel yang tidak relevan, serta mampu menangani dataset yang tidak terstruktur. Sementara itu, KNN dipilih karena kemudahan implementasi dan kemampuannya dalam menangani dataset dengan distribusi *non-linear.* Hasil penelitian menunjukkan bahwa *Random Forest* dengan *hyperparameter* memberikan akurasi tertinggi dan konsisten, menjadikannya pilihan yang lebih optimal dalam tugas klasifikasi.

Berdasarkan penelitian yang dilakukan, berikut adalah kesimpulan yang dapat diambil:

1. Hasil penelitian menunjukkan bahwa kedua metode klasifikasi tersebut memiliki tingkat akurasi yang tinggi, baik dengan menggunakan parameter *default* maupun setelah melakukan penalaan *hyperparameter*.
2. Dalam konteks penelitian ini, metode *Random Forest* dengan penalaan *hyperparameter* menghasilkan akurasi tertinggi dengan nilai 0.9807, serta memiliki AUC *Score* yang baik dengan nilai 0.9980. Oleh karena itu, metode ini dapat dijadikan pilihan yang lebih optimal dalam tugas penelitian klasifikasi untuk memprediksi COVID-19.

5.2 Saran

Untuk mendukung penelitian ini lebih lanjut tentang topik yang sama, ada beberapa saran sebagai berikut :

1. Pada penelitian ini hanya menggunakan dua metode saja sehingga perlu menambahkan metode data mining lainnya untuk mendapatkan perbandingan yang lebih banyak serta akurasi yang lebih besar
2. Perlu dilakukan uji hipotesis pada hasil berbandingan yang ada sehingga dapat diketahui seberapa besar tingkat signifikansinya