

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

1. Model *machine learning* berbasis *random forest* berhasil dibangun menggunakan *dataset* jaringan yang didapat dari *Kaggle*. Pembangunan model *machine learning* menggunakan beberapa parameter performa jaringan seperti *jitter*, *latency*, *throughput*, *packet loss*, *bandwith usage* dan 4 fitur hasil *feature engineering*. Proses pembangunan model melewati beberapa tahapan seperti pra-pemrosesan data, rekayasa fitur, normalisasi data, pelatihan model menggunakan *GridSearchCV*, dan kalibrasi probabilitas (*CalibratedClassifierCV*), serta *threshold tuning* untuk memaksimalkan *F1-score*.
2. Berdasarkan hasil yang didapatkan dari dua skema pembagian data, yaitu 60/20/20 dan 80/20. Pengukuran menunjukkan bahwa kedua model memiliki performa sangat tinggi. Model 60/20/20 dengan kalibrasi (*platt scaling*) dan *threshold tuning* mencapai akurasi *test* 99.5% dan *F1-score* 0.995, *precision* sebesar 1.00, serta *recall* yang mencapai 0.99. Model 80/20 dengan kalibrasi *sigmoid* (*cv=3*) menghasilkan akurasi *test* 100% dan *F1-score* 1.00, *precision* sebesar 1.00, serta *recall* 1.00. Validasi keseluruhan pada model 80/20 yang menggunakan *5-Fold Cross Validation* menunjukkan hasil *F1-score* rata-rata sebesar 98.42%, menandakan performa model yang dibangun stabil.

5.2 Saran

Dari hasil penelitian, serta analisis yang dilakukan pada penelitian ini, terdapat beberapa saran yang dapat diberikan dengan tujuan mengembangkan penelitian di masa mendatang dan juga saran yang bisa berguna untuk meningkatkan model lagi, berikut adalah beberapa saran:

1. Penelitian selanjutnya disarankan untuk menggunakan *dataset* yang baru atau melakukan evaluasi model menggunakan *dataset* dari sumber lain yang lebih beragam, atau data nyata dari jaringan operasional. Hal ini bertujuan untuk menguji kemampuan generalisasi model terhadap kondisi jaringan yang lebih kompleks.

2. Penggunaan simulasi jaringan seperti NS-3 dapat menjadi solusi untuk menghasilkan *dataset* yang lebih representatif dan realistis. Melalui simulasi, penelitian akan dapat mengontrol variasi *latency*, *packet loss*, dan *bandwidth* secara lebih akurat, serta menguji performa model dalam kondisi dinamis yang menyerupai dunia nyata.
3. Pada penelitian selanjutnya disarankan untuk menguji kembali model menggunakan *unseen dataset*, untuk menilai stabilitas dan mengidentifikasi potensi *overfitting* secara lebih komprehensif. Hal ini berhubungan dengan hasil akurasi, *recall*, *precision*, *F1-score* yang didapatkan pada penelitian ini mencapai akurasi yang sempurna, dan tidak adanya pengujian menggunakan *unseen dataset* karena pada penelitian ini *dataset* yang didapatkan untuk pengujian lebih lanjut tidak sesuai dengan parameter yang dibangun sehingga tidak bisa dilakukannya pengujian lebih lanjut, dan terdapat kendala waktu selama proses pembangunan.

