BAB V
PENUTUP

5.1 Kesimpulan

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

1. Penelitian ini berhasil menganalisis perilaku *malicious traffic* menggunakan log *honeypot* untuk mengidentifikasi indikator yang menggambarkan tingkat risiko serangan. Hasil analisis memperlihatkan bahwa fitur *connection_count, failed, success, unique_ports, unique_username, unique_password, session_count, failed_ratio,* dan *success_ratio* mampu merepresentasikan karakteristik aktivitas serangan yang terekam di *honeypot*. Selain itu, indikator *connection_count, failed, failed_ratio, unique_username*, dan *unique_password* berhasil dipakai sebagai dasar pembentukan *pseudo label* risiko menggunakan pendekatan *Threshold* Q90 dan *Risk Score* sehingga aktivitas bisa dikategorikan ke dalam kelas *Low Risk* dan *High Risk*.
2. Penelitian ini berhasil membangun dan mengevaluasi sistem klasifikasi risiko *malicious traffic* dengan algoritma *Random Forest* yang dapat mengelompokkan aktivitas serangan secara terukur dan sistematis ke dalam kategori *Low Risk* dan *High Risk*. Hasil evaluasi model memiliki nilai *accuracy* 99,40%, *precision* kelas *High Risk* sebesar 95,28%, *recall* sebesar 99,02%, *F1-score* sebesar 97,12%, dan *ROC-AUC* sebesar 99,95%. Selain itu, model yang sudah dilatih diimplementasikan dengan sukses ke dalam dashboard yang memiliki kemampuan untuk melakukan analisis otomatis terhadap data log *honeypot* dan menampilkan hasil klasifikasi risiko dalam bentuk statistik, tabel, grafik, dan rekomendasi keamanan. Dengan demikian, sistem yang dibangun dapat membantu proses analisis risiko *malicious traffic* secara lebih cepat dan berbasis data.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, maka saran yang dapat diberikan untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Penelitian selanjutnya disarankan menggunakan dataset yang telah memiliki label risiko aktual (*ground truth*) atau divalidasi oleh pakar keamanan siber

sehingga performa model dapat dievaluasi terhadap kondisi yang lebih mendekati lingkungan operasional nyata.

2. Penelitian selanjutnya dapat melakukan perbandingan algoritma *Random Forest* dengan metode *machine learning* lainnya seperti *Support Vector Machine* untuk memperoleh model dengan performa yang lebih optimal.
3. Sistem *dashboard* diharapkan dikembangkan agar mampu memproses *dataset* dengan struktur kolom yang berbeda melalui mekanisme *schema mapping* dan adaptasi fitur secara otomatis.
4. Pengembangan lebih lanjut dapat dilakukan dengan mengintegrasikan sistem ke dalam monitoring *real-time,* penambahan notifikasi otomatis, serta fitur rekomendasi mitigasi keamanan.
5. Sistem juga dapat dikembangkan dengan integrasi ke dalam platform keamanan seperti *SIEM* atau *IDS* agar dapat dimanfaatkan sebagai solusi yang lebih aplikatif dalam operasional keamanan siber.