BAB 5
SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini bertujuan mengevaluasi kinerja serta menganalisis pengaruh optimasi fitur terhadap dua pendekatan deteksi *Domain Generation Algorithm* (DGA), yaitu *Random Forest* dan *BiLSTM*. Dataset yang digunakan merupakan unifikasi dari berbagai sumber kredibel, meliputi UMUDGA, *Data Driven Security*, Kaggle, dan ExtraHop. Eksperimen meliputi pra-pemrosesan, pembagian data *hold-out* (80% latih dan 20% uji), penyeimbangan pada data latih melalui *undersampling*, pelatihan + validasi silang, serta evaluasi akhir pada data uji.

Berdasarkan hasil implementasi dan evaluasi yang telah dilakukan, diperoleh simpulan sebagai berikut (selaras dengan tujuan penelitian pada Bab 1):

1. Model BiLSTM murni (*Pure BiLSTM*) menunjukkan kinerja terbaik pada evaluasi data uji *hold-out* 20% dengan *accuracy* 0.9536, *precision* 0.9986, *recall* 0.9526, *F1-Score* 0.9751, dan ROC-AUC 0.9929. Hasil ini menegaskan bahwa pendekatan *deep learning* berbasis sekuensial lebih efektif dalam menangkap variasi pola DGA yang kompleks pada skala data besar.

2. Model *Random Forest* teroptimasi pada evaluasi data uji *hold-out* 20% menghasilkan *accuracy* 0.9195 dengan ROC-AUC 0.9888. Selain itu, model ini lebih baik dalam menekan *false positive* pada domain legit, tercermin dari *recall* kelas legit yang tinggi (0.9854). Dengan demikian, pendekatan berbasis fitur tetap relevan ketika prioritas utama adalah meminimalkan kesalahan pemblokiran pada domain legit.

3. Pada konfigurasi eksperimen yang digunakan, pendekatan BiLSTM *Hybrid* tidak meningkatkan metrik utama (*accuracy* dan *F1-score*) dibanding *Pure BiLSTM*. Model *Hybrid* menghasilkan *accuracy* 0.9497, *F1-Score* 0.9729, dan ROC-AUC 0.9932, sehingga penambahan fitur *handcrafted* pada arsitektur sekuensial pada penelitian ini cenderung belum memberikan peningkatan yang signifikan (indikasi adanya redundansi informasi terhadap representasi sekuens).

4. Analisis XAI menggunakan SHAP berhasil mengidentifikasi fitur dominan, seperti indikator kemiripan pola *n-gram* terhadap Tranco serta tingkat keacakan (misalnya entropi), yang mendorong prediksi ke kelas DGA atau legit. Hal ini meningkatkan transparansi keputusan model dan memudahkan verifikasi oleh analis keamanan.

Secara operasional, pemilihan model sebaiknya mengikuti prioritas risiko: jika fokus utama adalah performa agregat (*accuracy*/*F1-score*), maka *Pure BiLSTM* menjadi pilihan utama pada hasil penelitian ini, sedangkan jika fokus utama adalah menekan *false positive* (mengurangi risiko pemblokiran domain legit), maka *Random Forest* lebih sesuai. Untuk kebutuhan interpretabilitas, *Random Forest* + SHAP dapat digunakan sebagai pendamping analisis.

5.2 Saran

Berdasarkan hasil penelitian ini, beberapa saran untuk pengembangan lebih lanjut adalah sebagai berikut:

1. Melakukan eksplorasi arsitektur *deep learning* lain seperti Transformer atau kombinasi CNN-LSTM untuk menangkap pola yang lebih kompleks.
2. Mengembangkan mekanisme *online learning* agar model dapat beradaptasi terhadap famili DGA baru tanpa perlu pelatihan ulang dari awal.
3. Mengintegrasikan fitur kontekstual tambahan seperti informasi WHOIS, sertifikat SSL, atau pola lalu lintas DNS untuk meningkatkan akurasi deteksi.
4. Melakukan pengujian pada lingkungan operasional nyata untuk memvalidasi performa model dalam skenario *real-time*.