BAB V
KESIMPULAN

Bab ini menyajikan kesimpulan yang ditarik dari hasil analisis dan pembahasan pada bab sebelumnya, serta memberikan saran untuk penelitian selanjutnya. Kesimpulan dirumuskan untuk menjawab rumusan masalah mengenai pengaruh teknik seleksi fitur dan performa model RNN dalam mendeteksi serangan DDoS.

5.1. Kesimpulan

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, maka dapat disimpulkan beberapa hal berikut:

1. Penerapan teknik seleksi fitur *hybrid* yang menggabungkan *Learning Vector Quantization* (LVQ) dan *Autoencoder* terbukti memberikan dampak signifikan terhadap karakteristik model dalam mendeteksi serangan DDoS. Meskipun tidak menghasilkan akurasi tertinggi, pendekatan ini menawarkan trade-off antara performa mentah dan generalisasi model. Dampak utamanya dapat dirumuskan sebagai berikut:
   a. Efisiensi dan Stabilitas Model pada Proses seleksi fitur berhasil mereduksi dimensi data dari 85 menjadi 16 fitur terpilih. Reduksi ini secara langsung menyederhanakan arsitektur model dan mengurangi risiko *overfitting*. Hal ini terbukti pada kurva *loss*, di mana model *hybrid* menunjukkan tren penurunan yang lebih stabil dan seimbang dibandingkan model *baseline* yang turun drastis mendekati nol. Meskipun waktu pemrosesan total lebih lama (15.731 detik vs. 9.955 detik) karena adanya proses seleksi itu sendiri, model RNN yang dilatih dengan 16 fitur secara inheren lebih ringan dan tidak rentan "menghafal" data.
   b. Trade-off Performa pada Kelas Minoritas (BENIGN) pada Analisis performa pada kelas minoritas BENIGN menunjukkan adanya *trade-off* yang jelas antara metrik mentah yang superior dan generalisasi model yang lebih baik. **Model Baseline** mencatatkan metrik nyaris sempurna untuk kedua kelas, dengan **recall untuk kelas minoritas BENIGN mencapai 99,98%**. Performa ini, meskipun terlihat

superior, sangat dipengaruhi oleh dominasi kelas mayoritas ATTACK dan merupakan indikasi kuat adanya overfitting. **Model RNN_LVQ_FS** menjadi pembanding pertama yang menunjukkan pergeseran ke arah yang lebih seimbang. Model ini mampu **mempertahankan performa yang sangat tinggi**, dengan recall untuk kelas BENIGN hanya turun sedikit menjadi **99,88%**. Dengan presisi yang masih kuat di angka 90,3477%, LVQ terbukti menjadi kompromi terbaik yang mampu menjaga sensitivitas terhadap kelas minoritas sambil mengurangi kompleksitas fitur. **Model RNN_AE_FS dan Hybrid** menunjukkan *trade-off* yang jauh lebih signifikan. Model RNN_AE_FS mengalami penurunan recall BENIGN yang lebih tajam menjadi **92,06%** , dengan presisi anjlok ke 38,8733%. Sementara itu, model Hybrid berada di antara keduanya dengan recall BENIGN sebesar **94,89%**, namun dengan presisi terendah yaitu **20,9398%**, yang menyebabkan F1-Score yang rendah (34,2119%). Ini menunjukkan bahwa keuntungan dari model *hybrid* dan *AE* bukanlah peningkatan deteksi kelas minoritas karena performanya secara metrik berada di bawah *baseline* dan LVQ melainkan pencapaian model yang lebih general dan tidak terlalu bias pada pola kelas mayoritas.

c. Kombinasi Keunggulan Metode pada Pendekatan hybrid terbukti efektif karena mensintesis dua analitis yang saling melengkapi: Learning Vector Quantization (LVQ), sebagai metode *supervised*, bertindak pragmatis dengan memilih indikator serangan tingkat pertama yang paling jelas. Algoritma ini memprioritaskan fitur-fitur diskriminatif seperti Destination Port, Flow Duration, dan Fwd Packets/s untuk secara langsung **mengidentifikasi target dan volume serangan**. Autoencoder sebagai metode *unsupervised,* bertindak sebagai "analis struktural" yang ahli mendeteksi anomali. Algoritma mampu menangkap indikator serangan tingkat kedua yang lebih subtil, seperti gangguan pada pola temporal jaringan melalui fitur Inter Arrival Time (IAT) dan anomali perilaku melalui Down/Up

Ratio. Kombinasi kedua pendekatan ini menghasilkan sub-kumpulan fitur yang lebih informatif dan representatif. Hal ini memperkuat kemampuan model dalam mengenali pola serangan DDoS secara lebih menyeluruh, yang didukung oleh kinerja evaluasi yang lebih stabil dan kemampuan generalisasi yang lebih baik. Dibandingkan dengan model baseline yang cenderung "menghafal" pola mayoritas, pendekatan hybrid menunjukkan keunggulan dalam menangani data yang tidak seimbang, menghasilkan deteksi yang lebih adaptif dan dapat diandalkan untuk implementasi sistem keamanan jaringan di dunia nyata.

2. Performa model *Recurrent Neural Network* (RNN) berbasis LSTM sangat dipengaruhi oleh skenario seleksi fitur yang digunakan, yang menunjukkan adanya trade-off antara performa mentah dan generalisasi model. **Model** Baseline, yang tidak menggunakan seleksi fitur, mencatatkan metrik performa yang nyaris sempurna: akurasi 99,9997%, recall 99,9866%, dan presisi 99,8933%. *Confusion matrix* juga mengonfirmasi hal ini dengan kesalahan yang sangat minim (1 *False Negative*, 8 *False Positive*). Namun, performa superior ini tidak mencerminkan kemampuan generalisasi yang baik, melainkan lebih sebagai indikasi **overfitting** terhadap pola dominan dari kelas mayoritas ATTACK. Setelah dilakukan seleksi fitur dengan Learning Vector Quantization (LVQ), terjadi pergeseran ke arah performa yang lebih seimbang dan efisien. Akurasi tetap sangat tinggi di **99,9726%**. Model ini berhasil **mempertahankan sensitivitas yang sangat tinggi** terhadap kelas BENIGN, dengan *recall* untuk kelas tersebut hanya **turun sedikit** menjadi **99,88%** (dihitung dari 7.461 TP dan 9 FN). Penurunan presisi menjadi 90,3477% adalah *trade-off* yang wajar untuk mendapatkan model yang jauh lebih sederhana dan robust. Ini menunjukkan LVQ adalah **kompromi terbaik** antara performa dan efisiensi. Model yang menggunakan Autoencoder menunjukkan *trade-off* yang lebih signifikan. Akurasi turun menjadi 99,5375%, dan *recall* untuk kelas BENIGN juga turun lebih jauh ke **92,06%**. Anjloknya presisi ke 38,8733% karena tingginya *False Positive* (13.015) menunjukkan bahwa fitur hasil ekstraksi

unsupervised ini kurang optimal dalam memisahkan kedua kelas secara tegas. Terakhir, **Model Hybrid** memberikan hasil yang paling menonjolkan strategi pengutamaan generalisasi di atas metrik mentah. Meskipun akurasi turun ke 99,0542% dan presisi sangat rendah (20,9398%) , model ini masih menjaga *recall* untuk kelas BENIGN di angka **94,9%**. Ini menegaskan bahwa tujuan utama pendekatan seleksi fitur bukanlah untuk melampaui metrik *baseline* yang *overfit*, melainkan untuk menciptakan model yang lebih andal dan tidak terlalu bergantung pada pola kelas mayoritas. Kombinasi kedua pendekatan ini menghasilkan subset fitur yang lebih informatif dan representatif, yang memperkuat kemampuan model dalam mengenali pola serangan DDoS secara lebih menyeluruh. Hasil ini didukung oleh kinerja evaluasi yang lebih stabil dan kemampuan generalisasi yang lebih baik terhadap variasi data, sebagaimana tecermin pada kurva *loss* yang seimbang. Penurunan presisi yang terjadi pada model dengan seleksi fitur merupakan konsekuensi alami dari *trade-off* untuk mendapatkan model yang lebih general. Dalam konteks sistem deteksi DDoS, di mana keandalan dan pencegahan *overfitting* sangat krusial, pendekatan seleksi fitur seperti *hybrid* menawarkan solusi yang lebih adaptif. Dibandingkan dengan model *baseline* yang cenderung "menghafal" pola mayoritas, metode *hybrid* mendorong model agar tidak sekadar mengandalkan dominasi kelas dalam proses klasifikasi, menjadikannya lebih dapat diandalkan untuk implementasi sistem keamanan jaringan di dunia nyata. Oleh karena itu, pemilihan metode yang tepat harus disesuaikan dengan kondisi dan tujuan spesifik dari sistem deteksi yang akan dibangun Untuk sistem deteksi yang menuntut keseimbangan dan keandalan tinggi, di mana setiap alarm harus memiliki tingkat kepercayaan yang kuat dan jumlah alarm palsu harus minim, maka pendekatan **RNN dengan seleksi fitur LVQ** adalah pilihan yang paling unggul. Model ini menjaga performa deteksi nyaris sempurna sambil tetap efisien. Namun, untuk skenario sistem peringatan dini (*early warning*), di mana prioritas utamanya adalah **menangkap semua potensi serangan** agar tidak ada yang terlewat, maka pendekatan **RNN Hybrid** lebih sesuai . Meskipun presisinya sangat rendah dan menghasilkan banyak

alarm palsu, sensitivitasnya yang tinggi menjadikannya penyaring awal yang efektif untuk investigasi lebih lanjut

5.2. Saran

Berdasarkan hasil penelitian dan evaluasi yang telah dilakukan, serta mempertimbangkan tantangan yang ditemui selama proses pengembangan, berikut disampaikan beberapa saran untuk pengembangan dan penelitian lanjutan di masa mendatang:

1. Pengembangan Model Hybrid untuk Implementasi Nyata

   Model hybrid yang mengombinasikan seleksi fitur Learning Vector Quantization (LVQ) dan Autoencoder dalam arsitektur RNN berbasis LSTM menunjukkan performa paling seimbang, terutama dalam hal sensitivitas terhadap kelas minoritas dan efisiensi model. Oleh karena itu, disarankan agar pengembangan lebih lanjut difokuskan pada arsitektur ini untuk penerapan di lingkungan nyata, khususnya sistem deteksi intrusi berbasis real-time. Penelitian selanjutnya perlu mengkaji aspek latensi inferensi, efisiensi penggunaan sumber daya komputasi (memori, prosesor), serta daya tanggap model terhadap trafik jaringan langsung.

2. Eksplorasi Kombinasi Metode Seleksi Fitur yang Lebih Luas

   Keberhasilan pendekatan hybrid dalam penelitian ini menjadi indikator bahwa kombinasi teknik seleksi fitur dapat meningkatkan performa model secara signifikan. Penelitian lanjutan disarankan untuk mengeksplorasi kombinasi lain, seperti mengintegrasikan metode filter-based (misalnya: *Information Gain*, *Chi-Square*, *ReliefF*) dengan metode wrapper atau embedded (seperti *Recursive Feature Elimination* atau *Lasso Regression*). Tujuan utamanya adalah untuk menemukan konfigurasi fitur yang lebih optimal dalam menjaga keseimbangan antara akurasi, sensitivitas, dan kompleksitas model.

3. Optimasi Hyperparameter secara Sistematis

   Konfigurasi hyperparameter dalam penelitian ini masih bersifat manual dan terbatas. Untuk meningkatkan akurasi dan efisiensi model, disarankan melakukan proses tuning hyperparameter secara sistematis

dengan pendekatan seperti Grid Search, Random Search, atau Bayesian Optimization. Parameter penting yang perlu dievaluasi meliputi learning rate, jumlah unit LSTM, dropout rate, jumlah epoch, dan ukuran batch, agar model dapat lebih optimal terhadap karakteristik data yang digunakan.

4. Validasi Model pada Dataset Lain yang Lebih Variatif

   Agar model tidak hanya optimal pada dataset CICDDoS2019, disarankan melakukan uji validasi silang pada dataset DDoS lain yang memiliki karakteristik berbeda, seperti CICIDS2017, BoT-IoT, atau dataset terbaru dengan skenario serangan dan trafik jaringan yang lebih kompleks. Validasi silang ini dapat memperkuat argumen bahwa model memiliki daya generalisasi tinggi dan mampu mendeteksi serangan pada lingkungan yang berbeda.

5. Analisis Lebih Lanjut terhadap Kasus Kesalahan Klasifikasi (False Positive dan False Negative)

   Meskipun model hybrid menunjukkan performa terbaik secara global, tetap terdapat jumlah kesalahan klasifikasi yang signifikan, terutama dalam bentuk False Positive (trafik normal terdeteksi sebagai serangan). Penelitian lanjutan disarankan untuk melakukan analisis mendalam terhadap instance-instance yang salah diklasifikasikan, guna memahami pola-pola anomali atau noise yang menyebabkan model keliru. Hasil analisis ini dapat digunakan untuk merancang rekayasa fitur tambahan, atau untuk menyesuaikan struktur model agar lebih adaptif terhadap variasi trafik yang kompleks.

6. Perbaikan Urutan Proses Normalisasi dan Oversampling

   Dalam penelitian ini, normalisasi dilakukan sebelum pembagian data dan sebelum penerapan SMOTE, yang secara teknis dapat menyebabkan data leakage. Kondisi ini berisiko membuat informasi dari data uji ikut memengaruhi proses pelatihan model, sehingga hasil evaluasi menjadi bias. Ke depannya, disarankan agar proses normalisasi dilakukan hanya pada data latih, kemudian transformasi yang sama diterapkan pada data uji, agar validitas eksperimen tetap terjaga dan model tidak terpapar informasi dari data uji selama pelatihan.

7. Evaluasi Granular Berdasarkan Jenis Serangan DDoS

Evaluasi performa model dalam penelitian ini masih bersifat global tanpa mengidentifikasi efektivitas terhadap masing-masing jenis serangan DDoS, seperti SYN Flood, DNS Amplification, atau UDP Flooding. Penelitian selanjutnya disarankan untuk melakukan analisis performa per jenis serangan, agar diketahui kelemahan model terhadap tipe serangan tertentu. Informasi ini dapat dijadikan dasar untuk pengembangan sistem deteksi yang lebih spesifik dan responsif, terutama dalam lingkungan yang memiliki distribusi serangan yang dinamis.