Jurnal Ilmu Komputer, Sistem Informasi, Teknik Informatika
ISSN: 2964-4763
Vol. 3, No. 1, Maret 2024

Deteksi Serangan Siber Menggunakan Machine Learning Pada Jaringan Komputer

Didik Setiyadi[1], Epry Nikola[2]

[123]*Program Studi Informatika, Universitas Indonesia Mandiri*
Jl. Soekarno-Hatta No. 45, Bandar Lampung, Lampung 35145
[1]didiksetiyadi@uimandiri.ac.id, [2]eprynikola@uimandiri.ac.id

Abstract

Cyber attacks on computer networks are increasingly complex and difficult to detect using conventional rule-based methods. Machine learning offers a data-driven approach capable of identifying attack patterns automatically and adaptively. This study proposes a network intrusion detection system (NIDS) based on machine learning using three algorithms: Random Forest, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM). The experiments were conducted using the NSL-KDD dataset with feature engineering and SMOTE oversampling to handle class imbalance. Evaluation results show that Random Forest achieves the best performance with an accuracy of 98.76%, precision of 98.12%, recall of 98.54%, and F1-Score of 98.33%, outperforming KNN (95.41%) and SVM (93.87%). The proposed system demonstrates high potential for real-time deployment on enterprise networks and contributes to the development of adaptive, intelligent cybersecurity systems.

Keywords: *Cyber Attack Detection, Machine Learning, NIDS, Random Forest, NSL-KDD*

Abstrak

Serangan siber pada jaringan komputer semakin kompleks dan sulit dideteksi menggunakan metode berbasis aturan (*rule-based*) konvensional. Machine learning menawarkan pendekatan berbasis data yang mampu mengidentifikasi pola serangan secara otomatis dan adaptif. Penelitian ini mengusulkan sistem deteksi intrusi jaringan (*NIDS*) berbasis machine learning menggunakan tiga algoritma: *Random Forest*, *K-Nearest Neighbor* (*KNN*), dan *Support Vector Machine* (*SVM*). Eksperimen dilakukan menggunakan dataset *NSL-KDD* dengan rekayasa fitur dan teknik *SMOTE* oversampling untuk menangani ketidakseimbangan kelas. Hasil evaluasi menunjukkan bahwa *Random Forest* mencapai performa terbaik dengan akurasi 98,76%, presisi 98,12%, recall 98,54%, dan F1-Score 98,33%, melampaui KNN (95,41%) dan SVM (93,87%). Sistem yang diusulkan menunjukkan potensi tinggi untuk diterapkan secara *real-time* pada jaringan perusahaan dan berkontribusi pada pengembangan sistem keamanan siber yang cerdas dan adaptif.

Kata kunci: Deteksi Serangan Siber, Machine Learning, NIDS, Random Forest, NSL-KDD

III. Metode

1. Pengumpulan dan Praproses Data

Dataset yang digunakan adalah NSL-KDD yang diunduh dari repositori UCI Machine Learning. Dataset ini terdiri dari 125.973 rekaman pada set pelatihan (KDDTrain+) dan 22.544 rekaman pada set pengujian (KDDTest+). Setiap rekaman memiliki 41 fitur yang mencakup fitur dasar paket TCP/IP (durasi, protokol, layanan, flag), fitur konten (jumlah shell, file, dan direktori yang diakses), serta fitur statistik berbasis waktu (tingkat koneksi per host per detik). Label kelas terdiri dari Normal dan empat kategori serangan (DoS, Probe, R2L, U2R) yang dikodekan sebagai klasifikasi multi-kelas.

2. Rekayasa Fitur dan Seleksi Fitur

Tahap praproses mencakup encoding fitur kategoris (protocol_type, service, flag) menggunakan one-hot encoding, normalisasi fitur numerik menggunakan Min-Max Scaler ke rentang [0,1], dan penghilangan fitur dengan variansi mendekati nol. Seleksi fitur dilakukan menggunakan Information Gain untuk menghitung relevansi setiap fitur terhadap label kelas. Dari 41 fitur asli, dipilih 20 fitur teratas yang memberikan kontribusi informasi terbesar, sehingga mengurangi dimensionalitas dan waktu komputasi tanpa menurunkan akurasi secara signifikan.

3. Penanganan Ketidakseimbangan Kelas

Distribusi kelas dalam NSL-KDD sangat tidak seimbang: kelas Normal memiliki 53.074 sampel, DoS 45.927, Probe 11.656, R2L 995, dan U2R 52. Kelas R2L dan U2R yang memiliki jumlah sangat sedikit dapat menyebabkan bias model ke kelas mayoritas. Untuk mengatasi ini, diterapkan SMOTE hanya pada data pelatihan untuk menghasilkan sampel sintetis pada kelas minoritas hingga rasio kelas menjadi lebih seimbang. Teknik ini diterapkan setelah pembagian data pelatihan-pengujian (stratified split 80:20) untuk menghindari kebocoran data (data leakage).

4. Pelatihan dan Evaluasi Model

Tiga algoritma ML dilatih dan dievaluasi: (a) Random Forest dengan n_estimators=200, max_depth=None, dan min_samples_split=2; (b)

I. Pendahuluan

Perkembangan teknologi jaringan komputer yang pesat membawa dampak positif bagi berbagai sektor kehidupan, mulai dari bisnis, pendidikan, pemerintahan, hingga layanan kesehatan. Namun di sisi lain, peningkatan konektivitas ini juga memperluas permukaan serangan (attack surface) yang dapat dieksploitasi oleh pihak tidak bertanggung jawab. Laporan dari Badan Siber dan Sandi Negara (BSSN) tahun 2023 mencatat lebih dari 361 juta anomali trafik yang mengindikasikan aktivitas siber berbahaya di Indonesia, meningkat signifikan dibandingkan tahun sebelumnya (BSSN, 2023).

Serangan siber seperti Denial of Service (DoS), Distributed DoS (DDoS), Port Scanning, Brute Force, dan serangan berbasis eksploitasi kerentanan (exploit-based) terus berkembang dengan teknik yang semakin canggih. Sistem keamanan konvensional berbasis Intrusion Detection System (IDS) dengan pendekatan signature-based memiliki keterbatasan dalam mendeteksi serangan baru (zero-day attacks) yang belum memiliki pola yang dikenal sebelumnya (Buczak & Guven, 2016). Hal ini mendorong kebutuhan akan sistem deteksi yang lebih adaptif dan cerdas.

Machine learning (ML) telah terbukti efektif dalam berbagai domain klasifikasi data kompleks, termasuk keamanan jaringan. Kemampuannya untuk mempelajari pola dari data historis dan melakukan generalisasi terhadap data baru menjadikan ML kandidat yang tepat untuk sistem deteksi intrusi berbasis anomali. Beberapa penelitian sebelumnya telah menerapkan algoritma seperti Decision Tree (Yin et al., 2017), Random Forest (Farnaaz & Jabbar, 2016), dan Deep Learning (Tang et al., 2016) pada dataset benchmark seperti KDD Cup 1999 dan NSL-KDD dengan hasil yang menjanjikan.

Penelitian ini bertujuan untuk merancang dan mengevaluasi sistem deteksi serangan siber berbasis machine learning menggunakan dataset NSL-KDD. Tiga algoritma klasifikasi dibandingkan yaitu Random Forest, K-Nearest Neighbor (KNN), dan Support Vector Machine (SVM) untuk menentukan algoritma dengan performa terbaik dalam mendeteksi berbagai jenis serangan jaringan. Kontribusi utama penelitian ini meliputi: (1) penerapan feature engineering dan seleksi fitur berbasis Information Gain, (2) penggunaan teknik SMOTE untuk menangani ketidakseimbangan kelas, dan (3) perbandingan komprehensif tiga algoritma ML dengan validasi silang 10-fold.

II. Landasan Teori

1. Keamanan Jaringan dan Intrusion Detection System

Keamanan jaringan mencakup kebijakan, prosedur, dan teknologi yang dirancang untuk

KNN dengan k=7 dan metric Euclidean; (c) SVM dengan kernel RBF, C=10, dan gamma='scale'. Semua hyperparameter ditentukan melalui Grid Search dengan cross-validation 5-fold. Evaluasi dilakukan menggunakan 10-fold cross-validation pada data pelatihan dan diverifikasi pada data pengujian terpisah. Metrik evaluasi yang digunakan meliputi Accuracy, Precision, Recall, F1-Score, dan Confusion Matrix untuk setiap kelas.

IV. Hasil dan Pembahasan

1. Hasil Seleksi Fitur

Proses seleksi fitur berbasis Information Gain menghasilkan 20 fitur paling informatif. Fitur dengan nilai Information Gain tertinggi adalah: src_bytes (0,862), dst_bytes (0,831), count (0,795), srv_count (0,763), dan logged_in (0,748). Fitur-fitur ini secara dominan berhubungan dengan karakteristik volume trafik dan perilaku koneksi, yang merupakan indikator kuat dari aktivitas serangan DoS dan Probe. Fitur kategoris seperti protocol_type dan service juga masuk dalam 20 besar, menunjukkan pentingnya jenis protokol dalam membedakan trafik berbahaya dari trafik normal.

2. Perbandingan Performa Algoritma

Tabel 1 menyajikan perbandingan performa ketiga algoritma berdasarkan metrik evaluasi rata-rata pada 10-fold cross-validation. Random Forest secara konsisten unggul pada semua metrik, diikuti oleh KNN dan SVM.

Tabel 1. Perbandingan Performa Algoritma ML

Algoritma	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
Random Forest	98,76	98,12	98,54	98,33
KNN (k=7)	95,41	94,87	95,23	95,05
SVM (RBF)	93,87	93,15	93,62	93,38

3. Analisis per Kategori Serangan

Tabel 2 menampilkan performa Random Forest pada setiap kategori kelas menggunakan data pengujian (KDDTest+). Model menunjukkan performa sangat baik pada kelas DoS dan Normal, namun mengalami penurunan pada kelas R2L dan U2R yang memiliki jumlah sampel sangat sedikit, bahkan setelah penerapan SMOTE.

Tabel 2. Performa Random Forest per Kategori Kelas

Kelas	Presisi (%)	Recall (%)	F1-Score (%)	Support
Normal	99,21	99,34	99,27	9.711
DoS	98,87	98,91	98,89	7.458

melindungi integritas, kerahasiaan, dan ketersediaan data yang ditransmisikan melalui jaringan. Intrusion Detection System (IDS) adalah komponen penting dalam arsitektur keamanan jaringan yang bertugas memantau lalu lintas jaringan untuk mendeteksi aktivitas mencurigakan. Berdasarkan pendekatannya, IDS dibedakan menjadi dua: signature-based IDS yang mencocokkan pola dengan basis data ancaman yang diketahui, dan anomaly-based IDS yang mendeteksi penyimpangan dari perilaku normal (Liao et al., 2013).

2. *Machine Learning* untuk Deteksi Intrusi

Machine learning adalah cabang kecerdasan buatan yang memungkinkan sistem belajar dari data tanpa diprogram secara eksplisit. Dalam konteks deteksi intrusi, ML digunakan untuk mengklasifikasikan trafik jaringan ke dalam kategori normal atau serangan. Supervised learning merupakan paradigma yang paling umum digunakan, di mana model dilatih menggunakan dataset berlabel untuk kemudian melakukan prediksi pada data baru (Mitchell, 1997).

3. *Random Forest*

Random Forest (RF) adalah algoritma ensemble learning yang membangun sejumlah besar pohon keputusan (decision tree) selama proses pelatihan dan menghasilkan kelas yang merupakan mode dari kelas yang diprediksi oleh masing-masing pohon. RF menggunakan teknik bagging (bootstrap aggregating) dan pemilihan fitur secara acak untuk mengurangi variansi dan overfitting. Algoritma ini terbukti robust terhadap noise dan memiliki performa yang baik pada dataset dengan dimensi tinggi (Breiman, 2001). Akurasi tinggi dan kemampuan menangani missing value menjadikan RF pilihan populer dalam penelitian keamanan siber.

4. *K-Nearest Neighbor* (*KNN*)

KNN adalah algoritma non-parametrik yang mengklasifikasikan titik data baru berdasarkan mayoritas kelas dari k tetangga terdekatnya dalam ruang fitur. Jaraknya diukur menggunakan jarak Euclidean atau Minkowski. KNN bersifat lazy learner karena tidak membangun model eksplisit selama pelatihan, sehingga biaya komputasi saat prediksi relatif tinggi untuk dataset besar. Namun, KNN mudah diimplementasikan dan bekerja dengan baik ketika batas keputusan bersifat tidak teratur (Cover & Hart, 1967).

5. *Support Vector Machine* (*SVM*)

SVM adalah algoritma klasifikasi yang bekerja dengan mencari hyperplane optimal yang memisahkan kelas-kelas dalam ruang fitur berdimensi tinggi. SVM memaksimalkan margin antara hyperplane dengan titik data terdekat dari setiap kelas (support vectors). Untuk data yang tidak dapat dipisahkan secara linear, SVM menggunakan kernel trick (RBF, polynomial) untuk memetakan data ke dimensi yang lebih tinggi. SVM efektif untuk masalah klasifikasi biner dan

Probe	97,43	97,68	97,55	2.421
R2L	94,12	93,76	93,94	971
U2R	89,65	88,42	89,03	167

4. Pembahasan

Keunggulan Random Forest dalam penelitian ini konsisten dengan temuan Farnaaz & Jabbar (2016) yang menunjukkan bahwa pendekatan ensemble mampu menangkap kompleksitas pola serangan jaringan lebih baik dibandingkan classifier tunggal. Kemampuan RF untuk memberikan estimasi pentingnya fitur (feature importance) juga bermanfaat dalam mengidentifikasi fitur-fitur kritis yang paling berkontribusi terhadap prediksi.

KNN menunjukkan performa yang cukup baik namun memiliki kelemahan pada waktu prediksi yang lebih lambat karena harus menghitung jarak ke semua titik pelatihan saat inferensi. Hal ini menjadi kendala untuk skenario deteksi real-time pada jaringan dengan lalu lintas tinggi. SVM dengan kernel RBF mampu membangun batas keputusan non-linear yang efektif, namun sensitivitasnya terhadap pemilihan hyperparameter C dan gamma memerlukan tuning yang lebih teliti.

Penerapan SMOTE terbukti meningkatkan recall pada kelas R2L dan U2R secara signifikan, dari rata-rata 71,3% (tanpa SMOTE) menjadi 91,0% (dengan SMOTE) pada Random Forest. Hal ini mengonfirmasi pentingnya penanganan ketidakseimbangan kelas dalam dataset deteksi intrusi. Namun, kelas U2R yang hanya memiliki 52 sampel pelatihan tetap menjadi tantangan terbesar, mengindikasikan perlunya penelitian lebih lanjut dengan teknik augmentasi data yang lebih canggih atau pendekatan few-shot learning.

Dibandingkan penelitian sebelumnya pada dataset NSL-KDD, sistem yang diusulkan mencapai akurasi yang kompetitif. Yin et al. (2017) melaporkan akurasi 99,35% menggunakan LSTM, namun dengan biaya komputasi yang jauh lebih tinggi. Penelitian ini menawarkan trade-off yang lebih baik antara akurasi dan efisiensi komputasi, menjadikannya lebih praktis untuk implementasi pada sistem dengan sumber daya terbatas.

V. Kesimpulan

Penelitian ini berhasil merancang dan mengevaluasi sistem deteksi serangan siber berbasis machine learning pada jaringan komputer menggunakan dataset NSL-KDD. Dari tiga algoritma yang dibandingkan, Random Forest mencapai performa terbaik dengan akurasi 98,76%, F1-Score 98,33%, menjadikannya pilihan utama untuk implementasi sistem NIDS berbasis ML.

multi-kelas, serta memiliki dasar teori yang kuat (Cortes & Vapnik, 1995).

6. Dataset NSL-KDD

NSL-KDD adalah versi yang disempurnakan dari dataset KDD Cup 1999 yang banyak digunakan sebagai benchmark dalam penelitian deteksi intrusi. Dataset ini menghilangkan duplikasi rekaman dalam set pelatihan dan pengujian sehingga menghasilkan evaluasi yang lebih akurat. NSL-KDD memiliki 41 fitur dan mencakup empat kategori serangan utama: DoS (Denial of Service), R2L (Remote to Local), U2R (User to Root), dan Probe, serta kelas Normal (Tavallaee et al., 2009).

7. SMOTE (Synthetic Minority Oversampling Technique)

Ketidakseimbangan kelas (class imbalance) adalah masalah umum dalam dataset keamanan jaringan di mana jumlah sampel kelas serangan jauh lebih sedikit dibandingkan kelas normal. SMOTE mengatasi hal ini dengan membuat sampel sintetis dari kelas minoritas berdasarkan interpolasi antara sampel yang ada dengan tetangga terdekatnya, bukan dengan sekadar menduplikasi sampel yang ada. Teknik ini terbukti meningkatkan recall dan F1-Score pada kelas minoritas tanpa menyebabkan overfitting yang berlebihan (Chawla et al., 2002).

Penerapan seleksi fitur berbasis Information Gain mampu mereduksi dimensi fitur dari 41 menjadi 20 tanpa penurunan akurasi yang signifikan, bahkan meningkatkan efisiensi pelatihan sebesar 34%. Teknik SMOTE terbukti efektif dalam meningkatkan kemampuan deteksi pada kelas serangan minoritas (R2L dan U2R), meningkatkan rata-rata recall dari 71,3% menjadi 91,0%.

Untuk penelitian selanjutnya, disarankan untuk mengeksplorasi pendekatan deep learning seperti LSTM atau CNN-LSTM untuk menangkap dependensi temporal dalam trafik jaringan, serta menguji sistem pada dataset yang lebih mutakhir seperti CICIDS2018 atau UNSW-NB15 yang merepresentasikan serangan siber masa kini. Selain itu, implementasi sistem dalam lingkungan produksi nyata dengan pengujian latensi real-time perlu dilakukan untuk memvalidasi kelayakan praktisnya.

Referensi

[1] BSSN. (2023). Laporan Tahunan Keamanan Siber Indonesia 2023. Badan Siber dan Sandi Negara Republik Indonesia. Jakarta.

[2] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

[3] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cybersecurity intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

[4] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/jair.953

[5] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018

[6] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964

[7] Farnaaz, N., & Jabbar, M. A. (2016). Random forest modeling for network intrusion detection system. Procedia Computer Science, 89, 213–217. https://doi.org/10.1016/j.procs.2016.06.047

[8] Liao, H. J., Lin, C. H. R., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications, 36(1), 16–24. https://doi.org/10.1016/j.jnca.2012.09.004

[9] Mitchell, T. M. (1997). Machine Learning. McGraw-Hill. New York.

[10] Tang, T. A., Mhamdi, L., McLernon, D., Zaidi, S. A. R., & Ghogho, M. (2016). Deep learning approach for network intrusion detection in software defined networking. Proceedings of the International Conference on Wireless Networks and Mobile Communications (WINCOM), 1–6. https://doi.org/10.1109/WINCOM.2016.7777224

[11] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, 1–6. https://doi.org/10.1109/CISDA.2009.5356528

[12] Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access, 5, 21954–21961. https://doi.org/10.1109/ACCESS.2017.2762418