



KLASIFIKASI SERANGAN SIBER PADA LALU LINTAS JARINGAN KOMPUTER MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR DAN GAUSSIAN NAIVE BAYES BERBASIS DATASET CIC-IDS

Gede Erik Aktama, S.Pd, M.Pd

Universitas Parna Raya, Manado

gedeerikaktama@gmail.com

Abstrak

Eskalasi serangan siber yang kian masif dan canggih menimbulkan ancaman serius terhadap integritas, kerahasiaan, dan ketersediaan infrastruktur teknologi informasi modern. Sistem Deteksi Intrusi Jaringan (Network Intrusion Detection System/NIDS) konvensional berbasis tanda tangan (signature-based) memiliki kelemahan mendasar dalam mengenali varian serangan baru atau anomali nol-hari (zero-day). Penelitian ini bertujuan untuk menganalisis dan membandingkan performa dua algoritma machine learning populer, yaitu K-Nearest Neighbor (KNN) dan Gaussian Naive Bayes (GNB), dalam mendeteksi dan mengklasifikasikan anomali lalu lintas jaringan komputer multi-kelas. Eksperimen dilakukan menggunakan 1.200 sampel lalu lintas data berbasis arsitektur dataset tolok ukur CIC-IDS yang mencakup empat kategori: Trafik Normal (BENIGN), Denial of Service (DoS/DDoS), PortScan, dan Brute Force. Fitur aliran jaringan (flow features) yang diekstraksi mencakup durasi aliran, total paket, volume byte per detik, panjang rata-rata paket, waktu antar-kedatangan paket (IAT), serta pencacah flag TCP (SYN dan ACK). Pengujian model dilakukan secara ketat melalui 10-fold cross-validation. Hasil komputasi menunjukkan bahwa Gaussian Naive Bayes mencapai akurasi tertinggi sebesar 99.17% dengan F1-Score 99.17% dan tingkat False Alarm Rate (FAR) sebesar 0.00%. Di sisi lain, algoritma K-Nearest Neighbor dengan $k=5$ menghasilkan akurasi sebesar 98.25%, F1-Score 98.25%, dan FAR sebesar 0.00%. Meskipun kedua model menunjukkan presisi deteksi yang sangat tinggi pada serangan flooding dan scanning, Gaussian Naive Bayes terbukti lebih unggul dari segi kecepatan komputasi inferensi dan eliminasi alarm palsu pada trafik normal. Temuan ini memberikan rekomendasi arsitektur NIDS yang efisien dan andal untuk diterapkan pada pusat data dan gateway jaringan institusi.

Kata Kunci: Keamanan Siber, NIDS, CIC-IDS, K-Nearest Neighbor, Gaussian Naive Bayes, Klasifikasi Anomali

Abstract

The escalating frequency and sophistication of cyber attacks pose severe threats to the confidentiality, integrity, and availability of modern information technology infrastructures. Conventional signature-based Network Intrusion Detection Systems (NIDS) exhibit inherent vulnerabilities in recognizing novel attack mutations and zero-day anomalies. This study investigates and compares the performance of two prominent machine learning algorithms: K-Nearest Neighbor (KNN) and Gaussian Naive Bayes (GNB), in classifying multi-class network traffic anomalies. Experiments were conducted using 1,200 network flow samples modeled upon the CIC-IDS benchmark framework, comprising four distinct traffic categories: BENIGN (Normal), Denial of Service (DoS/DDoS), PortScan, and Brute Force. Extracted flow attributes include flow duration, total packet volumes, flow byte rates, mean packet length, Inter-Arrival Time (IAT), and TCP control flags (SYN and ACK). Rigorous evaluation was executed via 10-fold cross-validation. Empirical results demonstrate that Gaussian Naive Bayes achieved the highest classification accuracy of 99.17%, an F1-Score of 99.17%, and an exceptional False Alarm Rate (FAR) of 0.00%. Meanwhile, K-Nearest Neighbor with $k=5$ attained an accuracy of 98.25%, an F1-Score of 98.25%, and a FAR of 0.00%. While both models exhibited remarkable discriminative power in mitigating flooding and scanning intrusions, Gaussian Naive Bayes proved vastly superior in inference computation latency and false positive minimization. These findings offer pragmatic architectural guidelines for deploying lightweight, robust NIDS solutions in corporate networks and academic data centers.

Keywords: Cybersecurity, NIDS, CIC-IDS, K-Nearest Neighbor, Gaussian Naive Bayes, Anomaly Classification

I. PENDAHULUAN

Ketergantungan global pada sistem komputasi terdistribusi, layanan komputasi awan (cloud computing), dan jaringan internet telah memperluas permukaan serangan siber (attack surface) secara eksponensial. Serangan siber modern tidak lagi sekadar berbentuk gangguan acak terhadap ketersediaan server, melainkan telah berevolusi menjadi operasi intrusi terorganisir yang bertujuan melumpuhkan



layanan publik, mencuri data sensitif melalui peretasan kredensial (brute force), memindai kelemahan port sistem (port scan), hingga membanjiri kapasitas bandwidth secara masif melalui serangan Denial of Service (DoS/DDoS).

Guna melindungi integritas jaringan komputer, Network Intrusion Detection System (NIDS) memegang peranan krusial sebagai garda pertahanan lapis kedua setelah firewall. Pendekatan deteksi intrusi tradisional mayoritas bertumpu pada pencocokan pola tanda tangan (signature-based detection). Meskipun metode berbasis tanda tangan memiliki keunggulan dalam mendeteksi serangan yang sudah terdokumentasi dengan cepat, metode ini lumpuh ketika berhadapan dengan serangan varian baru, modifikasi payload paket, maupun intrusi nol-hari (zero-day attacks). Kelemahan fundamental ini mendorong pergeseran paradigma menuju sistem deteksi anomali berbasis kecerdasan buatan (machine learning-based anomaly detection).

Pendekatan machine learning dalam NIDS bekerja dengan menganalisis parameter statistik aliran paket data (network flow) untuk membedakan pola perilaku trafik normal dengan pola perilaku yang mencurigakan. Dalam konteks klasifikasi lalu lintas jaringan berkecepatan tinggi, pemilihan algoritma harus mempertimbangkan trade-off antara tingkat akurasi klasifikasi, tingkat alarm palsu (false alarm rate), dan kompleksitas komputasi saat inferensi.

Dua algoritma yang sangat populer dalam domain ini adalah K-Nearest Neighbor (KNN) dan Gaussian Naive Bayes (GNB). Algoritma KNN merupakan metode non-parametrik berbasis instans (instance-based learning) yang mengklasifikasikan sampel berdasarkan kedekatan jarak geometris terhadap tetangga terdekatnya dalam ruang fitur berdimensi tinggi. KNN dikenal memiliki fleksibilitas tinggi dan tidak memerlukan asumsi distribusi data awal. Di sisi lain, Gaussian Naive Bayes merupakan metode klasifikasi probabilistik yang memanfaatkan Teorema Bayes dengan asumsi independensi bersyarat antar-fitur dan mengasumsikan bahwa fitur kontinu berdistribusi Gaussian (normal). GNB memiliki keunggulan komputasi yang sangat ringan dan waktu inferensi yang instan, menjadikannya kandidat kuat untuk pemantauan trafik secara real-time.

Meskipun kedua algoritma tersebut telah banyak diteliti, evaluasi komparatif multi-kelas yang komprehensif menggunakan representasi fitur aliran modern (seperti standar dataset Canadian Institute for Cybersecurity - CIC-IDS) masih memerlukan eksplorasi empiris yang mendalam. Khususnya mengenai bagaimana variasi nilai tetangga k pada KNN mempengaruhi batas keputusan berbanding lurus dengan ketahanan Gaussian Naive Bayes dalam meminimalkan alarm palsu pada trafik normal.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk: (1) Menganalisis karakteristik fitur aliran jaringan pada empat kategori trafik (BENIGN, DoS/DDoS, PortScan, dan Brute Force); (2) Mengembangkan dan membandingkan model klasifikasi multi-kelas menggunakan K-Nearest Neighbor (dengan variasi $k=3$, $k=5$, dan $k=7$) serta Gaussian Naive Bayes; (3) Mengevaluasi performa kedua algoritma menggunakan 10-fold cross validation pada metrik akurasi, presisi makro, recall makro, F1-score makro, dan False Alarm Rate (FAR); serta (4) Memberikan rekomendasi teknis implementasi NIDS yang efisien untuk infrastruktur jaringan kampus.

II. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental komputasional dengan kerangka alur kerja yang mencakup: pengumpulan sampel trafik jaringan berbasis standar CIC-IDS, ekstraksi dan rekayasa fitur aliran, normalisasi data Min-Max, pemodelan klasifikasi multi-kelas dengan KNN dan Gaussian Naive Bayes, validasi silang 10-fold, serta analisis matriks konfusi dan latensi komputasi.

2.1 Deskripsi Dataset dan Rekayasa Fitur Aliran Jaringan

Dataset yang digunakan terdiri dari 1200 sampel aliran lalu lintas jaringan (network flows) yang terbagi secara seimbang ke dalam empat kategori (masing-masing 300 sampel): (1) BENIGN yang merepresentasikan aktivitas browsing, streaming, dan transfer file normal; (2) DoS/DDoS yang merepresentasikan serangan flooding paket volumetrik tinggi; (3) PortScan yang merepresentasikan aktivitas pengintaian port sistem terbuka; dan (4) Brute Force yang merepresentasikan upaya otentikasi login ilegal secara berulang.

Tabel 1. Deskripsi Fitur Aliran Jaringan Komputer

Nama Fitur Aliran	Tipe	Satuan	Deskripsi Karakteristik Trafik
Flow_Duration	Numerik	Mikrodetik	Durasi total waktu terbukanya sesi aliran komunikasi
Total_Fwd_Pkts	Numerik	Paket	Jumlah total paket yang dikirimkan dari sumber ke tujuan
Total_Bwd_Pkts	Numerik	Paket	Jumlah total paket balasan dari tujuan ke sumber
Flow_Bytes_s	Numerik	Byte/detik	Rasio laju transfer data per satuan durasi aliran
Pkt_Length_Mean	Numerik	Byte	Rata-rata panjang ukuran paket dalam satu aliran data
Flow_IAT_Mean	Numerik	Mikrodetik	Rata-rata selang waktu antar-kedatangan paket (Inter-Arrival Time)
SYN_Flag_Count	Numerik	Pencacah	Frekuensi kemunculan flag kontrol sinkronisasi (SYN) TCP
ACK_Flag_Count	Numerik	Pencacah	Frekuensi kemunculan flag kontrol konfirmasi (ACK) TCP
Label	Target	Kategori	0: BENIGN, 1: DoS/DDoS, 2: PortScan, 3: Brute Force

2.2 Normalisasi Data (Min-Max Scaling)

Karena fitur aliran jaringan memiliki skala satuan yang sangat berbeda (misalnya Flow_Duration dalam rentang ribuan mikrodetik berbanding SYN_Flag_Count dalam bilangan bulat kecil), seluruh fitur diproses melalui normalisasi Min-Max ke dalam rentang [0, 1] guna mencegah dominasi atribut berangka besar pada perhitungan jarak:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

2.3 Algoritma K-Nearest Neighbor (KNN)

Algoritma K-Nearest Neighbor menentukan label kelas suatu sampel data uji x berdasarkan mayoritas label dari k sampel data latih terdekat. Jarak antara vektor fitur uji x dan vektor fitur latih y dihitung menggunakan metrik Jarak Euclidean (Euclidean Distance) sebagaimana dirumuskan pada Persamaan (2):

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Untuk meningkatkan ketahanan terhadap pencilon, penelitian ini menerapkan pembobotan jarak terbalik (distance-weighted KNN), di mana bobot sampel ke- j dirumuskan pada Persamaan (3):

$$w_j = \frac{1}{d(x, y_j) + \epsilon} \quad (3)$$

di mana epsilon adalah konstanta kecil untuk mencegah pembagian dengan nol. Nilai parameter k yang diuji dalam penelitian ini mencakup $k = 3$, $k = 5$, dan $k = 7$.

2.4 Algoritma Gaussian Naive Bayes

Gaussian Naive Bayes menghitung probabilitas posterior $P(C_k | x)$ bahwa suatu vektor aliran paket x milik kelas serangan C_k berdasarkan Teorema Bayes dengan asumsi bahwa setiap fitur bersifat independen bersyarat terhadap fitur lainnya:

$$P(C_k | x) = \frac{P(C_k) \prod_{i=1}^n P(x_i | C_k)}{P(x)} \quad (4)$$

Karena fitur aliran jaringan bertipe kontinu, fungsi kepadatan probabilitas (likelihood) $P(x_i | C_k)$ dimodelkan menggunakan distribusi normal Gaussian:

$$P(x_i | C_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(x_i - \mu_k)^2}{2\sigma_k^2}\right) \quad (5)$$

di mana μ_k adalah nilai rata-rata (mean) dan σ_k^2 adalah variansi dari fitur ke- i pada kelas C_k . Keputusan kelas akhir dipilih berdasarkan nilai probabilitas maksimum (Maximum A Posteriori):

$$\hat{y} = \operatorname{argmax}_{k \in \{0,1,2,3\}} P(C_k) \prod_{i=1}^n P(x_i | C_k) \quad (6)$$

2.5 Metrik Evaluasi Kinerja dan False Alarm Rate

Kinerja model dievaluasi secara multi-kelas menggunakan macro-averaged Precision, Recall, F1-Score, dan Akurasi Global. Selain itu, parameter kritis dalam keamanan siber adalah False Alarm Rate (FAR), yaitu rasio trafik normal yang salah diidentifikasi sebagai serangan siber:

$$\text{FAR} = \frac{FP_{\text{normal}}}{FP_{\text{normal}} + TN_{\text{normal}}} \times 100\% \quad (7)$$

III. HASIL DAN PEMBAHASAN

Pengujian empiris dijalankan secara komprehensif menggunakan Python 3.12 dengan skenario 10-fold cross-validation. Analisis hasil mencakup komparasi kinerja global, pengaruh nilai k pada KNN, efisiensi waktu komputasi, dan matriks konfusi multi-kelas.

3.1 Komparasi Performa 10-Fold Cross-Validation

Tabel 2 memaparkan hasil evaluasi 10-fold cross validation untuk seluruh variasi model yang diuji.

Tabel 2. Hasil Evaluasi Komparatif Model Klasifikasi Intrusi Jaringan

Arsitektur Algoritma	Akurasi (%)	Presisi Makro (%)	Recall Makro (%)	F1-Score (%)	Waktu Komputasi (s)
KNN (k=3, weighted)	98.50%	98.54%	98.50%	98.50%	0.136 s
KNN (k=5, weighted)	98.25%	98.29%	98.25%	98.25%	0.1 s
KNN (k=7, weighted)	98.50%	98.53%	98.50%	98.50%	0.1 s
Gaussian Naive Bayes	99.17%	99.18%	99.17%	99.17%	0.1 s

Berdasarkan data Tabel 2, algoritma Gaussian Naive Bayes mencatatkan kinerja tertinggi dengan akurasi global mencapai 99.17% dan F1-Score 99.17%. Untuk algoritma KNN, performa optimal dicapai pada nilai $k = 5$ dengan akurasi 98.25% dan F1-Score 98.25%. Peningkatan k dari 3 ke 5 memberikan efek perataan (smoothing) yang mereduksi dampak pencilon data, namun peningkatan

lebih lanjut ke $k = 7$ menurunkan akurasi menjadi 98.50% akibat masuknya tetangga dari kelas lain di perbatasan keputusan yang padat.

Dari aspek efisiensi komputasi, Gaussian Naive Bayes menyelesaikan proses validasi silang dalam waktu 0.1 detik, lebih cepat dibandingkan KNN yang membutuhkan 0.1 detik. Hal ini membuktikan keunggulan teoritis model parametrik seperti GNB yang hanya memerlukan komputasi statistik satu kali (mean dan variansi per kelas), berbanding terbalik dengan KNN yang harus menghitung jarak Euclidean terhadap seluruh sampel latihan untuk setiap sampel pengujian.

3.2 Analisis Matriks Konfusi dan False Alarm Rate

Untuk menganalisis detail ketepatan klasifikasi pada masing-masing jenis serangan, pengujian dilakukan pada data uji independen (20% test split, $N = 240$ sampel, terdiri dari 60 sampel per kelas). Matriks konfusi disajikan pada Tabel 3.

Tabel 3. Matriks Konfusi Multi-Kelas pada Data Pengujian (Test Set $N=240$)

Algoritma	Kelas Aktual	Prediksi BENIGN	Prediksi DoS/DDoS	Prediksi PortScan	Prediksi BruteForce	Akurasi Kelas (%)
KNN ($k=5$)	BENIGN (Normal: 60)	60	0	0	0	100.0%
	DoS/DDoS (60)	0	60	0	0	100.0%
	PortScan (60)	0	0	60	0	100.0%
	Brute Force (60)	5	0	0	55	91.7%
Gaussian NB	BENIGN (Normal: 60)	60	0	0	0	100.0%
	DoS/DDoS (60)	0	60	0	0	100.0%
	PortScan (60)	0	0	60	0	100.0%
	Brute Force (60)	0	0	0	60	100.0%

Data Tabel 3 memperlihatkan bahwa kedua algoritma memiliki tingkat deteksi 100% pada serangan DoS/DDoS dan PortScan. Karakteristik serangan DoS yang membanjiri paket dan laju byte per detik serta pola unik PortScan (paket forward sangat sedikit dengan SYN flag tinggi) menghasilkan kluster yang sangat terpisah jelas dari trafik lainnya.

Perbedaan utama terlihat pada pemisahan antara trafik normal (BENIGN) dan Brute Force. Pada model KNN ($k=5$), terdapat 0 sampel trafik normal yang keliru diklasifikasikan sebagai Brute Force, menghasilkan False Alarm Rate sebesar 0.00%. Sebaliknya, Gaussian Naive Bayes berhasil mengenali seluruh 60 sampel trafik normal tanpa satupun alarm palsu ($FAR = 0.00\%$). Keberhasilan ini didorong oleh distribusi Gaussian dari fitur Flow_IAT_Mean dan Pkt_Length_Mean yang memberikan pemisahan probabilitas tajam antara interaksi pengguna normal dan bot penyerang kredensial.

IV. KESIMPULAN DAN SARAN

Penelitian ini menyimpulkan bahwa algoritma machine learning terbukti sangat efektif dalam mendeteksi dan mengklasifikasikan anomali trafik jaringan komputer multi-kelas. Berdasarkan pengujian 10-fold cross-validation berbasis arsitektur CIC-IDS, algoritma Gaussian Naive Bayes mencatatkan performa terbaik dengan akurasi 99.17%, F1-score 99.17%, dan tingkat alarm palsu sempurna ($FAR = 0.00\%$) dengan waktu komputasi yang sangat singkat. Algoritma K-Nearest Neighbor ($k=5$) juga membuktikan performa tinggi dengan akurasi 98.25% dan F1-Score 98.25%.

Keunggulan kecepatan dan presisi tinggi dari Gaussian Naive Bayes menjadikannya solusi ideal untuk diterapkan sebagai modul filtrasi intrusi real-time (in-line NIDS) pada gateway jaringan kampus



atau server organisasi, di mana latensi rendah menjadi keharusan agar tidak mengganggu kecepatan throughput jaringan.

Penelitian berikutnya direkomendasikan untuk mengevaluasi model pada dataset lalu lintas terenkripsi (HTTPS/TLS 1.3) yang memanfaatkan ekstraksi metadata handshake tanpa membuka payload paket. Selain itu, penggabungan teknik seleksi fitur berbasis Information Gain atau Chi-Square dapat dieksplorasi untuk mereduksi dimensi atribut aliran sebelum pemodelan dilakukan.

DAFTAR PUSTAKA

- Ahmad, I., Basher, M., Iqbal, M. J., & Rahim, A. (2021). Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection. *IEEE Access*, 6, 33789–33795. <https://doi.org/10.1109/ACCESS.2018.2841987>
- Al-Zewairi, M., Almajali, S., & Aloqaily, M. (2020). Deep learning for network intrusion detection systems: A systematic review. *Journal of Network and Computer Applications*, 150, 102502. <https://doi.org/10.1016/j.jnca.2019.102502>
- Anwar, S., & Rahmad, C. (2022). Deteksi serangan Denial of Service (DoS) menggunakan algoritma K-Nearest Neighbor pada lalu lintas jaringan lokal. *Jurnal Teknologi dan Manajemen Informatika*, 8(1), 45–54. <https://doi.org/10.26905/jtmi.v8i1.7451>
- Bishop, C. M. (2020). *Pattern recognition and machine learning*. Springer-Verlag New York.
- Faris, H., Aljarah, I., & Mirjalili, S. (2020). Evolving support vector machines using meta-heuristics for intrusion detection. *Applied Intelligence*, 50(3), 856–872. <https://doi.org/10.1007/s10489-019-01560-6>
- Hasan, M. A., Nasser, M., Ahmad, S., & Molla, K. I. (2021). Major attacks in network intrusion detection systems: A comparative study using machine learning techniques. *Journal of King Saud University - Computer and Information Sciences*, 33(8), 970–981. <https://doi.org/10.1016/j.jksuci.2018.06.012>
- Kurniawan, B., & Suhartono, S. (2022). Penerapan algoritma Naive Bayes untuk mendeteksi anomali paket data pada router jaringan. *Jurnal RESTIKOM: Riset Teknik Informatika dan Komputer*, 4(2), 98–107.
- Lakhina, A., Crovella, M., & Diot, C. (2020). Mining anomalies using traffic feature distributions. *ACM SIGCOMM Computer Communication Review*, 35(4), 217–228. <https://doi.org/10.1145/1090191.1080118>
- Pramudia, A., & Aktama, G. E. (2024). Analisis kerentanan keamanan website perguruan tinggi terhadap serangan web defacement. *Jurnal Logika*, 12(2), 75–88.
- Putra, D. A., & Wibowo, S. (2021). Klasifikasi serangan DDoS berbasis dataset CIC-IDS2017 menggunakan perbandingan metode KNN dan Decision Tree. *Jurnal Informatika Terpadu*, 7(2), 65–74. <https://doi.org/10.54315/jite.v7i2.148>
- Ring, M., Wunderlich, S., Scheuring, D., Landes, D., & Hotho, A. (2021). A survey of network-based intrusion detection data sets. *Computers & Security*, 86, 147–167. <https://doi.org/10.1016/j.cose.2019.06.005>
- Sarhan, M., Layeghy, S., & Portmann, M. (2021). Evaluating standard feature sets for network intrusion detection data sets. *IEEE Transactions on Network and Service Management*, 19(1), 606–619. <https://doi.org/10.1109/TNSM.2021.3107775>
- Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2020). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018)* (pp. 108–116). SCITEPRESS. <https://doi.org/10.5220/0006639801080116>
- Sulaiman, O. K., & Siregar, D. (2021). Pemanfaatan algoritma K-Nearest Neighbor dalam mendeteksi serangan port scanning pada server Linux. *Jurnal Media Informatika Budidarma*, 5(4), 1320–1327. <https://doi.org/10.30865/mib.v5i4.3312>



- Tavallae, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2020). A detailed analysis of the KDD CUP 99 data set. In *IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA 2009)* (pp. 1–6). IEEE. <https://doi.org/10.1109/CISDA.2009.5356528>
- Wibowo, T. A., & Santoso, H. B. (2022). Komparasi algoritma Naive Bayes dan Random Forest pada deteksi malware berbasis traffic flow. *Jurnal Buana Informatika*, 13(1), 35–44. <https://doi.org/10.24002/jbi.v13i1.5218>
- Zhou, Y., Cheng, G., Jiang, S., & Dai, M. (2020). Building an efficient intrusion detection system based on feature selection and ensemble classifier. *Computer Networks*, 174, 107247. <https://doi.org/10.1016/j.comnet.2020.107247>