



Volume 4, No 1, October (2026)	DOI https://doi.org/10.59585/jimad	Page 48-55
-----------------------------------	--	---------------

Klasifikasi Multi-Kelas Serangan Jaringan Menggunakan Algoritma Random Forest pada Dataset NSL-KDD

Ajeng Hidayati*

*Program Studi Informatika, Universitas Pertahanan Republik Indonesia

***Corresponding Author:** Ajeng Hidayati; Email: ajeng.hidayati@idu.ac.id

ARTICLE INFO

Keywords: Machine Learning, Klasifikasi Serangan Jaringan, Random Forest, Intrusion Detection System, NSL-KDD.

Received : September 2026

Revised : September 2026

Accepted : September 2026

ABSTRACT

Cyberattacks on computer networks are increasing in both volume and complexity, necessitating intrusion detection systems capable of accurately classifying attack types. This study aims to compare the performance of Random Forest, Decision Tree, and Naive Bayes algorithms in classifying network traffic into five categories: Normal, Denial of Service (DoS), Probe, Remote to Local (R2L), and User to Root (U2R). The methodology involves a multi-class classification experiment using the NSL-KDD benchmark dataset, comprising data preprocessing, model training, and evaluation based on accuracy, precision, recall, and F1-score metrics on independent test data (KDDTest+). The results indicate that the Decision Tree achieved the highest accuracy (76.15%), slightly outperforming Random Forest (75.00%), while Naive Bayes lagged significantly behind (28.71%). Both decision tree-based models effectively detected Normal and DoS categories but struggled significantly to detect R2L and U2R attacks due to class imbalance in the training data and the presence of novel attack patterns in the test data. Feature importance analysis identified **src_bytes**, **same_srv_rate**, and **flag** as the most influential features. These findings underscore the need for data balancing strategies to improve the detection of minority attack classes in machine learning-based network security systems.



Volume 4, No 1, October (2026)	DOI https://doi.org/10.59585/jimad	Page 48-55
-----------------------------------	--	---------------

PENDAHULUAN

Pertumbuhan infrastruktur jaringan komputer dan internet yang pesat turut memperluas permukaan serangan (attack surface) yang dapat dieksploitasi oleh pihak yang tidak bertanggung jawab. Berbagai bentuk serangan siber, mulai dari Denial of Service (DoS), pemindaian jaringan (Probe), hingga upaya eskalasi hak akses (User to Root/U2R) dan akses jarak jauh tanpa otorisasi (Remote to Local/R2L), terus berkembang mengikuti evolusi teknologi jaringan itu sendiri. Mekanisme keamanan konvensional berbasis signature, seperti firewall dan antivirus tradisional, memiliki keterbatasan dalam mendeteksi pola serangan baru karena bergantung pada basis data tanda tangan serangan yang bersifat statis (Fernandes, Rodrigues, & Carvalho, 2019).

Sistem deteksi intrusi (Intrusion Detection System/IDS) berbasis machine learning menjadi pendekatan alternatif yang menjanjikan karena mampu mempelajari pola anomali dari data lalu lintas jaringan tanpa bergantung sepenuhnya pada basis data signature. Berbagai algoritma klasifikasi seperti Random Forest, Decision Tree, Support Vector Machine, dan Naive Bayes telah banyak diterapkan dan diuji pada dataset benchmark keamanan jaringan, salah satunya NSL-KDD (Negandhi, Trivedi, & Mangrulkar, 2019), sebagaimana dirangkum secara komprehensif dalam survei Liu dan Lang (2019). Pendekatan berbasis pohon keputusan juga terus dikembangkan melalui skema hibrida, misalnya kombinasi decision tree dengan support vector machine dan clustering hierarkis (Zou, Luo, Zhang, Yang, & Wang, 2023), serta dikaji secara mendalam dalam tinjauan komparatif terhadap berbagai algoritma IDS (Azam, Islam, & Huda, 2023). Dataset ini dikembangkan sebagai perbaikan dari dataset KDD Cup 1999 yang memiliki masalah redundansi data signifikan, sehingga cenderung membuat model klasifikasi menjadi bias terhadap rekaman yang sering muncul (Tavallaee, Bagheri, Lu, & Ghorbani, 2009).

Meskipun banyak penelitian melaporkan akurasi klasifikasi yang sangat tinggi, mendekati 99%, pada dataset NSL-KDD, sebagian besar hasil tersebut diperoleh dari evaluasi pada data validasi yang berasal dari distribusi sama dengan data latih. Ketika model diuji pada data uji independen (KDDTest+) yang secara sengaja memuat proporsi serangan baru yang tidak muncul pada data latih, penurunan akurasi yang signifikan kerap terjadi. Sebagai contoh, penelitian Wu, Fan, Zhu, You, Zhou, dan Huang (2022) melaporkan akurasi hingga 99,72% pada data latih, namun turun menjadi hanya 78,47% ketika diuji pada KDDTest+. Fenomena ini menunjukkan adanya kesenjangan antara kinerja model pada kondisi ideal dan kinerja model pada kondisi pengujian yang lebih realistis, sekaligus menyoroti tantangan generalisasi model IDS terhadap serangan yang belum pernah dipelajari sebelumnya. Selain itu, masalah ketidakseimbangan kelas (class imbalance) pada dataset NSL-KDD, di mana kategori serangan R2L dan U2R memiliki jumlah sampel jauh lebih sedikit dibandingkan kategori Normal dan DoS, turut menjadi tantangan tersendiri dalam membangun model klasifikasi yang mampu mendeteksi seluruh jenis serangan secara merata (Wu et al., 2022).

Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk: (1) membandingkan kinerja tiga algoritma klasifikasi, yaitu Random Forest, Decision Tree, dan Naive Bayes, dalam mengklasifikasikan lalu lintas jaringan ke dalam lima kategori pada dataset NSL-KDD menggunakan skema evaluasi train-test independen yang realistis; (2) menganalisis fitur-fitur jaringan yang paling berpengaruh terhadap keputusan klasifikasi melalui analisis feature importance; dan (3) mengidentifikasi tantangan spesifik dalam mendeteksi kategori serangan minoritas (R2L dan U2R) sebagai dasar rekomendasi penelitian lanjutan.



Volume 4, No 1, October (2026)	DOI https://doi.org/10.59585/jimad	Page 48-55
-----------------------------------	--	---------------

METODE

Penelitian ini menggunakan pendekatan eksperimental kuantitatif dengan skema klasifikasi multi-kelas (multi-class classification) terhadap lalu lintas jaringan. Dataset yang digunakan adalah NSL-KDD, yang terdiri atas data latih (KDDTrain+) sebanyak 125.973 rekaman dan data uji independen (KDDTest+) sebanyak 22.544 rekaman. Setiap rekaman memiliki 41 fitur numerik dan kategorikal yang menggambarkan karakteristik koneksi jaringan (seperti durasi koneksi, protokol, jumlah byte terkirim/diterima, dan berbagai rasio statistik lalu lintas), serta satu label yang menunjukkan jenis aktivitas, baik normal maupun salah satu dari 22 jenis serangan spesifik (Tavallae et al., 2009).

Label serangan yang bersifat spesifik, misalnya neptune, smurf, satan, guess_passwd, dan buffer_overflow, dipetakan ke dalam lima kategori utama sesuai konvensi standar pada literatur NSL-KDD, yaitu Normal, DoS (Denial of Service), Probe, R2L (Remote to Local), dan U2R (User to Root). Fitur kategorikal (protocol_type, service, flag) dikonversi menjadi nilai numerik menggunakan Label Encoding, sedangkan kolom difficulty level tidak digunakan dalam proses klasifikasi.

Tiga algoritma klasifikasi diterapkan dan dibandingkan kinerjanya. Pertama, Random Forest, yaitu algoritma ensemble learning berbasis kumpulan pohon keputusan yang dilatih pada subset data secara acak (bootstrap aggregating) dengan jumlah estimator sebanyak 100 pohon. Kedua, Decision Tree, yaitu algoritma pohon keputusan tunggal berbasis pemisahan rekursif menggunakan kriteria Gini index. Ketiga, Naive Bayes, yaitu algoritma klasifikasi probabilistik berbasis Teorema Bayes dengan asumsi independensi antar fitur (Gaussian Naive Bayes), yang telah lama digunakan sebagai baseline pada berbagai penelitian deteksi intrusi (Panda & Patra, 2007; Koc, Mazzuchi, & Sarkani, 2012). Ketiga model dilatih menggunakan data KDDTrain+ dan dievaluasi pada data KDDTest+ yang independen dan tidak pernah dilihat model selama pelatihan, guna mensimulasikan kondisi deteksi pada lalu lintas jaringan yang belum diketahui sebelumnya.

Kinerja setiap model dievaluasi menggunakan empat metrik standar, yaitu akurasi (accuracy), presisi (precision), recall, dan F1-score, dengan skema pembobotan weighted average untuk mengakomodasi ketidakseimbangan jumlah sampel antar kelas. Confusion matrix disusun untuk menganalisis pola kesalahan klasifikasi antar kategori secara lebih rinci. Selain itu, dilakukan analisis feature importance pada model Random Forest untuk mengidentifikasi sepuluh fitur jaringan yang paling berkontribusi terhadap keputusan klasifikasi, mengingat pemilihan dan pemahaman fitur relevan terbukti memengaruhi kinerja klasifikasi IDS secara signifikan (Thakkar & Lohiya, 2021). Seluruh eksperimen diimplementasikan menggunakan bahasa pemrograman Python dengan pustaka scikit-learn (Pedregosa et al., 2011).

HASIL DAN PEMBAHASAN

1. Hasil

a. Distribusi Data Hasil Praproses

Setelah dilakukan pemetaan label serangan spesifik ke dalam lima kategori utama, diperoleh distribusi data sebagaimana ditampilkan pada Tabel 1. Kategori Normal dan DoS mendominasi baik pada data latih maupun data uji, sementara kategori R2L dan terutama U2R memiliki jumlah sampel yang sangat sedikit. Ketidakseimbangan ini menjadi salah satu faktor utama yang memengaruhi kemampuan model dalam mendeteksi kedua kategori tersebut, sebagaimana akan dibahas lebih lanjut.

**Tabel 1.** Distribusi Kategori Serangan pada Dataset NSL-KDD

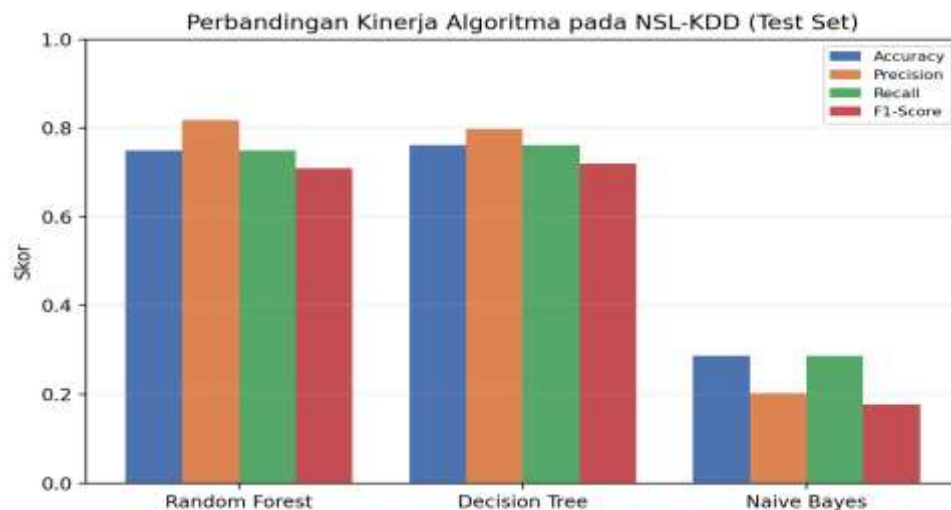
Kategori	Jumlah Data Latih	Jumlah Data Uji
Normal	67.343	9.711
DoS	45.927	7.460
Probe	11.656	2.421
R2L	995	2.885
U2R	52	67
Total	125.973	22.544

b. Perbandingan Kinerja Antar Algoritma

Ketiga algoritma dievaluasi pada data uji independen KDDTest+ menggunakan empat metrik utama. Hasil perbandingan kinerja secara keseluruhan (weighted average) ditampilkan pada Tabel 2 dan Gambar 1.

Tabel 2. Perbandingan Kinerja Algoritma pada Data Uji (KDDTest+)

Algoritma	Akurasi	Presisi	Recall	F1-Score	Waktu Latih (detik)
Random Forest	75,00%	81,77%	75,00%	70,88%	7,90
Decision Tree	76,15%	79,78%	76,15%	71,94%	0,81
Naive Bayes	28,71%	20,20%	28,71%	17,71%	0,29

**Gambar 1.** Perbandingan Kinerja Algoritma pada NSL-KDD (Test Set)

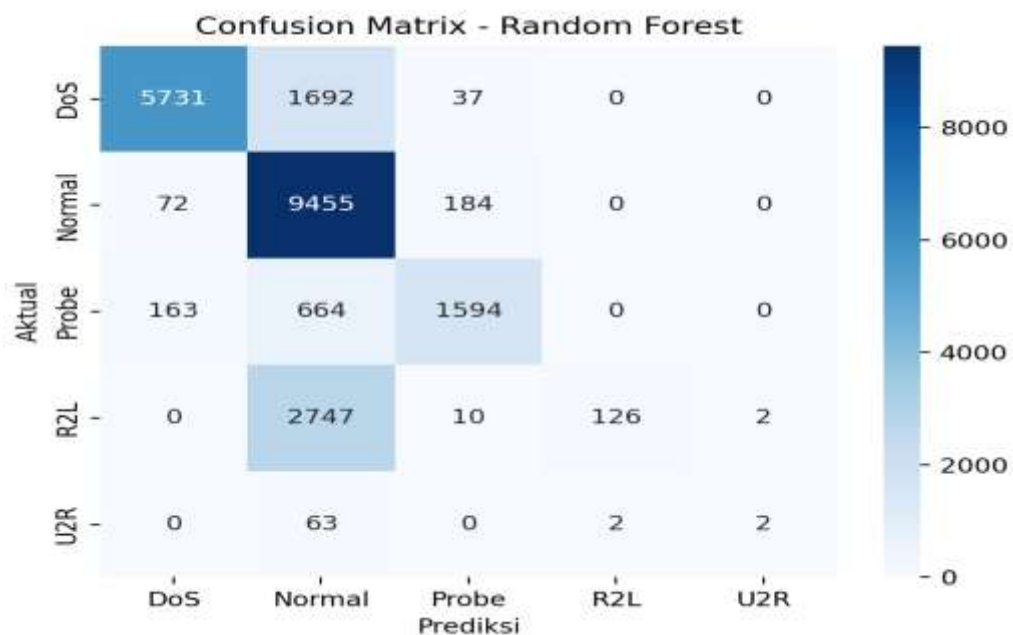


Decision Tree memperoleh akurasi tertinggi (76,15%), sedikit unggul dibandingkan Random Forest (75,00%), meskipun Random Forest justru memiliki presisi weighted yang lebih tinggi (81,77% berbanding 79,78%). Hal ini mengindikasikan bahwa Random Forest cenderung lebih konservatif dalam memprediksi kelas minoritas, sehingga prediksi yang dihasilkan lebih presisi namun recall keseluruhannya sedikit lebih rendah. Naive Bayes menunjukkan kinerja yang jauh lebih rendah pada seluruh metrik, konsisten dengan karakteristik algoritma tersebut yang mengasumsikan independensi antar fitur, sebuah asumsi yang tidak realistis untuk data lalu lintas jaringan karena banyak fitur NSL-KDD (misalnya `same_srv_rate`, `dst_host_serror_rate`, dan `srv_serror_rate`) secara inheren saling berkorelasi. Keunggulan relatif Decision Tree pada penelitian ini juga sejalan dengan temuan Chen, Song, Zhang, Li, dan Yang (2023), yang melaporkan bahwa model berbasis decision tree yang dioptimalkan mampu memberikan akurasi, recall, dan F-score yang kompetitif pada tugas klasifikasi intrusi jaringan menggunakan NSL-KDD.

Dari sisi efisiensi komputasi, Decision Tree dan Naive Bayes jauh lebih cepat dilatih dibandingkan Random Forest, karena Random Forest membangun 100 pohon keputusan secara ensemble. Meskipun demikian, selisih waktu latih ini (kurang dari 8 detik pada 125.973 data) masih tergolong sangat efisien untuk kebutuhan praktis sistem deteksi intrusi berbasis batch.

c. Analisis Confusion Matrix dan Kinerja per Kategori

Untuk memahami pola kesalahan klasifikasi secara lebih rinci, confusion matrix model Random Forest ditampilkan pada Gambar 2, sedangkan kinerja per kategori (precision, recall, F1-score) untuk seluruh algoritma ditampilkan pada Tabel 3.



Gambar 2. Confusion Matrix Model Random Forest pada KDDTest+

**Tabel 3.** Precision, Recall, dan F1-Score per Kategori (RF = Random Forest, DT = Decision Tree)

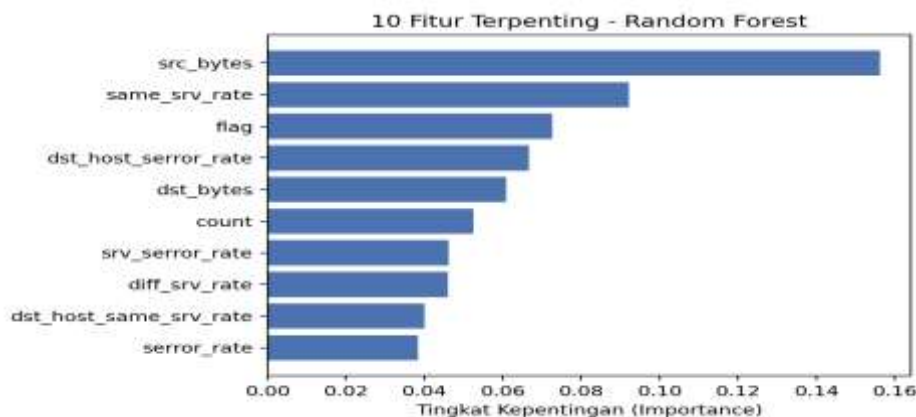
Kategori	Precision	Recall	F1 RF	Precision	Recall	F1 DT
	RF	RF		DT	DT	
Normal	64,7%	97,4%	77,7%	68,4%	96,6%	80,1%
DoS	96,1%	76,8%	85,4%	95,2%	82,8%	88,6%
Probe	87,3%	65,8%	75,1%	67,3%	59,6%	63,2%
R2L	98,4%	4,4%	8,4%	89,3%	5,2%	9,8%
U2R	50,0%	3,0%	5,6%	54,2%	19,4%	28,6%

Confusion matrix pada Gambar 2 memperlihatkan bahwa sebagian besar kesalahan klasifikasi Random Forest terjadi pada kategori R2L, di mana 2.747 dari 2.885 sampel R2L (95,2%) justru diprediksi sebagai Normal. Pola serupa, meskipun tidak sedrastis itu, juga terjadi pada kategori U2R dan Probe. Sebaliknya, kategori Normal dan DoS menunjukkan precision dan recall yang jauh lebih baik pada kedua model berbasis pohon keputusan.

Temuan ini konsisten dengan hasil penelitian Wu et al. (2022) yang juga melaporkan F1-score sangat rendah untuk kategori U2R dan R2L (di bawah 10%) meskipun akurasi keseluruhan model mereka mencapai 78,47% pada KDDTest+. Rendahnya recall pada R2L dan U2R umumnya disebabkan oleh dua faktor utama: (1) jumlah sampel latih yang sangat sedikit (995 dan 52 sampel) membuat model kesulitan mempelajari pola serangan tersebut secara memadai, dan (2) NSL-KDD Test+ secara sengaja memuat variasi jenis serangan R2L dan U2R yang tidak muncul sama sekali pada data latih, sebagai bagian dari desain dataset untuk menguji kemampuan generalisasi model terhadap serangan baru (Tavallae et al., 2009). Tantangan serupa dalam mendeteksi kelas minoritas pada NSL-KDD juga dilaporkan oleh Iftikhar, Rehman, Shah, Alenazi, dan Ali (2025), yang menunjukkan bahwa peningkatan F1-score pada kategori DoS dan Probe jauh lebih mudah dicapai dibandingkan kategori serangan dengan jumlah sampel latih yang sangat terbatas.

d. Analisis Feature Importance

Analisis feature importance pada model Random Forest, sebagaimana ditampilkan pada Gambar 3, mengidentifikasi sepuluh fitur jaringan yang paling berkontribusi terhadap keputusan klasifikasi.

**Gambar 3.** Sepuluh Fitur Terpenting Berdasarkan Random Forest



Volume 4, No 1, October (2026)	DOI https://doi.org/10.59585/jimad	Page 48-55
-----------------------------------	--	---------------

Fitur `src_bytes` (jumlah byte data yang dikirim dari sumber ke tujuan) memiliki tingkat kepentingan tertinggi (15,63%), diikuti oleh `same_srv_rate` (9,21%) dan `flag` (7,27%). Dominasi `src_bytes` sejalan dengan karakteristik serangan DoS dan Probe yang umumnya menghasilkan pola volume data atau jumlah koneksi yang menyimpang signifikan dari lalu lintas normal. Fitur-fitur berbasis rasio koneksi (`same_srv_rate`, `diff_srv_rate`, `serror_rate`, `srv_error_rate`) juga berkontribusi besar, mengindikasikan bahwa pola perilaku koneksi berulang dalam rentang waktu tertentu merupakan penciri kuat aktivitas anomali dibandingkan atribut konten paket itu sendiri.

Temuan ini memiliki implikasi praktis: sistem IDS berbasis machine learning dapat difokuskan pada pemantauan fitur-fitur statistik lalu lintas (traffic-based features) alih-alih hanya menganalisis payload paket data, sehingga berpotensi lebih efisien secara komputasi untuk implementasi deteksi secara real-time.

KESIMPULAN

Penelitian ini membandingkan kinerja tiga algoritma machine learning, yaitu Random Forest, Decision Tree, dan Naive Bayes, dalam mengklasifikasikan lalu lintas jaringan ke dalam lima kategori pada dataset NSL-KDD. Hasil eksperimen menunjukkan bahwa Decision Tree memperoleh akurasi tertinggi (76,15%), diikuti Random Forest (75,00%), sementara Naive Bayes memiliki kinerja jauh lebih rendah (28,71%) akibat asumsi independensi antar fitur yang kurang sesuai dengan karakteristik data lalu lintas jaringan yang saling berkorelasi.

Kedua model berbasis pohon keputusan mampu mendeteksi kategori Normal dan DoS dengan baik, namun mengalami kesulitan signifikan dalam mendeteksi kategori R2L (recall di bawah 6%) dan U2R (recall di bawah 20%) akibat ketidakseimbangan jumlah sampel pada data latih serta adanya pola serangan baru pada data uji yang tidak dipelajari model. Analisis feature importance mengidentifikasi `src_bytes`, `same_srv_rate`, dan `flag` sebagai fitur jaringan yang paling berpengaruh terhadap keputusan klasifikasi.

Temuan ini menegaskan bahwa evaluasi model IDS pada data uji independen memberikan gambaran kinerja yang lebih realistis dibandingkan evaluasi pada data dengan distribusi yang sama dengan data latih. Untuk penelitian selanjutnya, disarankan penerapan teknik penyeimbangan data seperti SMOTE (Chawla, Bowyer, Hall, & Kegelmeyer, 2002), eksplorasi metode ensemble lanjutan, serta feature selection untuk meningkatkan kemampuan deteksi terhadap kategori serangan minoritas, khususnya R2L dan U2R yang secara nyata merepresentasikan ancaman keamanan siber yang serius meskipun jarang terjadi.

REFERENSI

- Azam, Z., Islam, M. M., & Huda, M. N. (2023). Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. *IEEE Access*, 11, 80348–80391. <https://doi.org/10.1109/ACCESS.2023.3296444>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- 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>



Volume 4, No 1, October (2026)	DOI https://doi.org/10.59585/jimad	Page 48-55
-----------------------------------	--	---------------

- Chen, D., Song, Q., Zhang, Y., Li, L., & Yang, Z. (2023). Identification of network traffic intrusion using decision tree. *Journal of Sensors*, 2023, Article 5997304. <https://doi.org/10.1155/2023/5997304>
- Fernandes, G., Rodrigues, J. J. P. C., & Carvalho, L. F. (2019). A comprehensive survey on network anomaly detection. *Telecommunication Systems*, 70(3), 447–489. <https://doi.org/10.1007/s11235-018-0475-8>
- Iftikhar, N., Rehman, M. U., Shah, M. A., Alenazi, M. J. F., & Ali, J. (2025). Intrusion detection in NSL-KDD dataset using hybrid self-organizing map model. *Computer Modeling in Engineering & Sciences*, 143(1), 639–671. <https://doi.org/10.32604/cmes.2025.062788>
- Koc, L., Mazzuchi, T. A., & Sarkani, S. (2012). A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. *Expert Systems with Applications*, 39(18), 13492–13500.
- Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. *Applied Sciences*, 9(20), 4396. <https://doi.org/10.3390/app9204396>
- Negandhi, P., Trivedi, Y., & Mangrulkar, R. (2019). Intrusion detection system using random forest on the NSL-KDD dataset. In N. Shetty, L. M. Patnaik, H. C. Nagaraj, P. N. Hamsavath, & N. Nalini (Eds.), *Emerging research in computing, information, communication and applications* (pp. 519–529). Springer. https://doi.org/10.1007/978-981-13-6001-5_43
- Panda, M., & Patra, M. R. (2007). Network intrusion detection using naive bayes. *International Journal of Computer Science and Network Security*, 7(12), 258–263.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- 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 (CISDA)*. <https://doi.org/10.1109/CISDA.2009.5356528>
- Thakkar, A., & Lohiya, R. (2021). Attack classification using feature selection techniques: a comparative study. *Journal of Ambient Intelligence and Humanized Computing*, 12(1), 1249–1266. <https://doi.org/10.1007/s12652-020-02167-9>
- Wu, T., Fan, H., Zhu, H., You, C., Zhou, H., & Huang, X. (2022). Intrusion detection system combined enhanced random forest with SMOTE algorithm. *EURASIP Journal on Advances in Signal Processing*, 2022, Article 39. <https://doi.org/10.1186/s13634-022-00871-6>
- Zou, L., Luo, X., Zhang, Y., Yang, X., & Wang, X. (2023). HC-DTTSVM: A network intrusion detection method based on decision tree twin support vector machine and hierarchical clustering. *IEEE Access*, 11, 21404–21416. <https://doi.org/10.1109/ACCESS.2023.3251354>