JITE (Journal of Informatics and Telecommunication Engineering)

Available online <http://ojs.uma.ac.id/index.php/jite> DOI : 10.31289/jite.v5i2.6112

Received: 15 November 2021

Accepted: 29 December 2021

Published: 26 January 2022

Impact of Feature Selection Methods on Machine Learning-based for Detecting DDoS Attacks : Literature Review

Muhammad Nur Faiz1)*, Oman Somantri2), Abdul Rohman Supriyono3), Arif Wirawan Muhammad4)

1), 2),3)Teknik Informatika, Politeknik Negeri Cilacap, Indonesia

4) Teknik Informatika, Fakultas Informatika, IT Telkom Purwokerto, Indonesia

*Corresponding Email faiz@pnc.ac.id

Abstrak

Serangan keamanan siber menjadi semakin canggih dan meningkat dengan perkembangan teknologi sehingga menghadirkan ancaman baik bagi sektor swasta atau publik, terutama serangan Denial of Service (DoS) dan variannya yang sering dikenal dengan Distributed Denial of Service (DDoS). Cara untuk meminimalisir serangan ini salah satunya dengan menggunakan solusi mitigasi tradisional seperti Teknik analisis trafik jaringan dengan bantuan manusia, akan tetapi mengalami beberapa batasan dan masalah kinerja. Untuk mengatasi keterbatasan ini, Machine Learning (ML) telah menjadi salah satu teknik utama untuk memperkaya, melengkapi, dan meningkatkan pengalaman keamanan tradisional. Cara kerja ML ini didasarkan pada proses pengumpulan data, pelatihan dan output. ML ini dipengaruhi oleh beberapa faktor, salah satunya adalah rekayasa fitur. Pada penelitian ini, berfokus pada kajian literatur beberapa penelitian terbaru yang menunjukkan bahwa proses seleksi fitur sangat berdampak dengan tingkat akurasi dari ML ini. Dataset seperti KDD, UNSW-NB15 dan lainnya juga berpengaruh ada tingkat akurasi dari ML. Berdasarkan kajian literatur ini, Penelitian ini dapat mengamati beberapa strategi rekayasa fitur dengan dampak yang relevan dapat dipilih untuk meningkatkan solusi ML pada serangan DDoS. Kata Kunci: Aplikasi, pelayanan, web

Kata Kunci: DDos, Machine Learning, Dataset, Seleksi Fitur.

Abstract

Cybersecurity attacks are becoming increasingly sophisticated and increasing with the development of technology so that they present threats to both the private and public sectors, especially Denial of Service (DoS) attacks and their variants which are often known as Distributed Denial of Service (DDoS). One way to minimize this attack is by using traditional mitigation solutions such as human-assisted network traffic analysis techniques but experiencing some limitations and performance problems. To overcome these limitations, Machine Learning (ML) has become one of the main techniques to enrich, complement and enhance the traditional security experience. The way ML works are based on the process of data collection, training and output. ML is influenced by several factors, one of which is feature engineering. In this study, we focus on the literature review of several recent studies which show that the feature selection process greatly impacts the level of accuracy of this ML. Datasets such as KDD, UNSW-NB15 and others also affect the level of accuracy of ML. Based on this literature review, this study can observe several feature engineering strategies with relevant impacts that can be chosen to improve ML solutions on DDoS attacks.

Keywords: DDos, Machine Learning, Datasets, Feature engineering.

How to Cite: Faiz, M. N., Somantri, O., Supriyono, A. R., & Muhammad, A. W. (2022). Impact of Feature Selection Methods on Machine Learning-based for Detecting DDoS Attacks : Literature Review. *JITE (Journal of Informatics and Telecommunication Engineering)*, 5(2), 305–314.

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi (TIK) yang semakin pesat, aktivitas sehari-hari masyarakat kini semakin banyak dilakukan dengan menggunakan internet menyebabkan jumlah lalu lintas jaringan meningkat. Karena TIK telah menjadi bagian penting dari kehidupan modern, banyak jenis

serangan dan variasinya, termasuk penolakan layanan (DoS), serangan man-in-the-middle, dan sniffer, serta malware. (Haider et al., 2020) (Parmar & Gosai, 2015). Salah satu jenis dari DoS adalah DDoS (Distributed Denial of Service). Serangan DDoS ini bertujuan untuk mengganggu layanan dengan mengirimkan paket yang melebihi kapasitas mesin sumber yang ditargetkan untuk menanggapi permintaan. Pada tahun 2020, DDoS (distributed denial-of-service) attack mengalami peningkatan dibandingkan tahun sebelumnya. NETSCOUT misalnya mengobservasi sebanyak 10.089.687 DDoS attack pada tahun 2020; lebih tinggi dari sekitar 8.500.000 pada tahun 2019. Salah satu “senjata” yang digunakan penyerang dalam melakukan DDoS attack adalah perangkat IoT (internet of things). NETSCOUT pun menyebutkan contoh malware yang secara spesifik menasar perangkat IoT dan beredar di dunia meningkat lebih dari 100% pada tahun 2020 dibandingkan sebelumnya. DDoS attack ini dipercaya pula akan tetap populer pada tahun 2021 (Hildebrand, 2021).

Teknik DDoS ini memanfaatkan penyerang menggunakan komputer zombie yang dibuat dengan perangkat lunak berbahaya yang ditempatkan di komputer korban untuk mengirim paket dalam jumlah besar. Serangan DDoS menyebabkan lalu lintas jaringan yang tinggi dengan paket yang dikirim di jaringan, menyebabkan pengguna biasa yang ingin menerima layanan tidak menanggapi permintaan mereka. Serangan DDoS merupakan teknik yang sangat populer digunakan oleh hacker untuk mengambil alih sebuah server. Sampai saat ini serangan DDoS masih termasuk dalam jenis ancaman utama keamanan cyber. Deteksi dini memegang peran fundamental dalam mencegah dampak fatal serangan DDoS terhadap sumberdaya server. Salah satu tindakan dasar yang dilaksanakan untuk mencegah serangan DDoS adalah memasang Intrusion Detection System (IDS) pada server untuk memonitor aliran paket data yang masuk dalam jaringan internal ataupun sebaliknya (Fuertes et al., 2017). Teknik deteksi pada common IDS/IPS jauh dari sempurna apabila dibandingkan dengan variasi teknik dan tools modern yang digunakan oleh attacker karena IDS masih menggunakan model signature-based detection ataupun anomaly-based detection (Chowdhury, 2017). Penggunaan model deteksi yang ada pada IDS berbasis signature maupun anomaly menimbulkan rerata false-positive yang tinggi. Ditinjau dari segi teknis, IDS signature-based dan IDS anomaly-based bekerja dengan cara melakukan monitoring terhadap aliran paket data yang masuk ataupun yang keluar dari jaringan internal. IDS akan memberikan penanda apabila ia menemukan aktivitas aliran data yang tidak sesuai dengan database signature yang telah tertanam didalam IDS (Ambusaidi et al., 2016).

Model deteksi tersebut secara logis akan menimbulkan banyak false positive flag, karena aliran paket data jaringan komputer memiliki sifat dinamis baik dari segi ukuran, sumber, protokol, maupun konten isi data (Zhang et al., 2016). Di sisi lain, IDS signature-based dan IDS anomaly-based memiliki dua kelemahan utama. Kelemahan yang pertama adalah ketika IDS dalam mendeteksi serangan yang dimulai dengan protokol SYN misalnya SYN-Flood, disebabkan karena protokol SYN adalah protokol legal dan mutlak untuk digunakan untuk memulai komunikasi antara dua buah komputer/device dalam jaringan (Ramadhan et al., 2016). Oleh sebab itu, ordinary IDS sulit untuk membangkitkan alert terhadap serangan yang dimulai dengan artefak protokol SYN. Kelemahan yang kedua dari IDS terutama disebabkan oleh defisit protokol TCP/IP yang menyebabkan mudahnya attacker untuk memulai serangan DDoS misalnya dengan menggunakan perintah Ping yang secara default tersedia di seluruh sistem operasi ataupun menggunakan tools khusus semisal HOIC, LOIC, XOIC, golden-eye, dan sebagainya (Machaka & Bagula, 2016). Penggunaan protokol standar TCP/IP oleh attacker untuk melaksanakan serangan DDoS menyebabkan lambatnya target untuk menyadari bahwa ia sedang dibawah serangan, sehingga berdampak pula pada lambatnya proses mitigasi serangan (Hajar et al., 2016). Kelemahan protokol TCP/IP itulah yang sulit untuk dihandle oleh ordinary IDS. Selain itu, tingginya volume false positive flag yang dihasilkan oleh ordinary IDS cukup berdampak pada efisiensi hardening server. Berdasarkan kelemahan yang ada pada ordinary IDS tersebut, penelitian ini memiliki tujuan untuk mendeteksi serangan DDoS dengan mengutilisasi teknik feature engineering dan machine learning sehingga bisa menjadi improvement dalam pengembangan perangkat IDS.

Literatur review ini dilakukan untuk mendapatkan suatu teori yang bisa mendukung pemecahan suatu masalah yang akan atau sedang diteliti. Teori ini dapat dijadikan langkah pertama peneliti agar dapat lebih memahami alur kerangka berpikir ilmiah. Literature review ini juga dapat dijadikan sebagai pembanding beberapa penelitian sebelumnya sehingga nantinya dapat menghasilkan kebaruan dari penelitian yang dilakukan.

II. METODE

Penelitian ini berfokus pada melakukan kajian literatur dari beberapa penelitian sebelumnya yang masih berhubungan dengan topik yang akan diteliti. Langkah pertama yang dilakukan dalam membuat literatur review adalah mencari literatur yang relevan dengan penelitian, sehingga akan membantu dalam mendapatkan gambaran dari suatu topik penelitian. Kemudian, melihat dari literatur yang ada, apa saja yang menjadi kontribusi tentang topik yang dibahas. Langkah selanjutnya adalah meringkas literatur yang sudah ada. Setelah itu, peneliti melakukan pembahasan secara deskriptif berdasarkan penelitian yang sebelumnya. Langkah ini akan bermanfaat bagi akademisi yang ingin melakukan penelitian tentang minat perilaku dan perilaku pengguna terhadap penerimaan suatu teknologi informasi.

1. Studi DDoS

Serangan DDoS ini memiliki banyak bentuk seperti SYN banjir, banjir UDP, banjir ICMP dan sebagainya. Beberapa bentuk penanggulangan serangan ini seperti pemberian bandwidth, update software, firewall, software antivirus, dan Intrusion Detection System (IDS), dll (Wirawan et al., 2020). Menurut (Jaafar et al., 2019) ada 4 fase yang terlibat dalam pertahanan serangan DDoS adalah:

A. Pencegahan

Fase ini berfokus pada perlindungan sistem dan server dari serangan eksternal. Serangan DDoS ini biasanya dijalankan oleh sistem secara otomatis, mungkin dapat diatasi dengan membuat situs web yang hanya mengizinkan pengguna nyata untuk mengakses layanan dan sumber daya server web. Desain web harus efisien sehingga dapat menunda penyerang untuk masuk (Jaafar et al., 2019) (Mahjabin et al., 2017) (Merouane, 2017).

B. Mitigasi

Fase ini menangani serangan yang masuk dan memperlambatnya. Mitigasi juga untuk menghentikan serangan, sehingga teknik ini memblokir alamat IP dari lalu lintas jaringan yang tidak sah. Dampaknya adalah untuk menjamin pengguna otentik yang dapat mengakses layanan web (Mahjabin et al., 2017) (Bârli et al., 2021).

C. Deteksi

Sebuah fase yang membutuhkan analisis mendalam untuk menemukan anomali dalam lalu lintas jaringan. Deteksi ini melibatkan data pelatihan seperti pencocokan pola, klasifikasi, statistik, analisis deviasi, asosiasi dan lain-lain (Sonule et al., 2020) (Feroz Khan & Anandharaj, 2020).

D. Pemantauan

Fase ini untuk memonitor secara real-time untuk mendeteksi serangan DDoS. Ini bisa menjadi sulit jika penyerang memanfaatkan botnet yang terletak di beberapa lokasi di seluruh dunia (Novaes et al., 2021) (Marta et al., 2020).

Gambar 1. Fase pertahanan serangan DDoS (Jaafar et al., 2019) (Renuga Devi & Umamaheswari, 2020)

Sistem Deteksi instruksi adalah solusi yang umum digunakan untuk menganalisis dan mendeteksi serangan DDoS. Teknik IDS ini memiliki tujuan untuk mendeteksi, mengklasifikasi dan merespon serangan DDoS yang mempengaruhi integritas, kerahasiaan atau ketersediaan sumber daya jaringan (Min et al.,

2018)(Thakkar & Lohiya, 2021). Sistem ini didasarkan pada dua metode pendeteksian, yaitu Deteksi Penyalahgunaan dan Deteksi Anomali (Cepheli et al., 2016). Deteksi Penyalahgunaan juga dikenal sebagai Deteksi tanda tangan, Deteksi pola, Deteksi berbasis pengetahuan atau deteksi aturan (Thapngam et al., 2014). Teknik deteksi ini menyaring tanda tangan sesuai dengan database (Lima Filho et al., 2019). Ini lebih efisien dalam memperkenalkan pola paket, tetapi untuk mendeteksi bentuk ancaman dan varian baru, pertama-tama harus diperkenalkan ke system (Moustafa & Slay, 2016). Deteksi Anomali Teknik pendeteksian ini lebih berfokus pada trafik jaringan (Torabi et al., 2021). Keuntungan dari teknik ini adalah lebih cepat untuk mengetahui serangan baru tetapi tidak memberikan informasi yang jelas tentang dampak dari serangan tersebut. Teknik ini juga dapat mempengaruhi efisiensi dan kecepatan dalam mendeteksi intrusi (Zhou et al., 2017) (Kunhare & Tiwari, 2018).

2. Studi Machine Learning (ML)

Machine Learning (ML) adalah bidang Artificial Intelligence (AI) yang berisi aturan, metode dan fungsi untuk menangani suatu masalah (Marvi et al., 2020). Dalam DDoS, teknik ML ini belajar dari DDoS dan dataset normal, sehingga teknik ini mengetahui mana yang aman dan mana yang tidak. ML ini sangat bergantung pada pelatihan sebelumnya dengan memanfaatkan kumpulan data yang dikumpulkan. Sebuah model ML memiliki dua bagian: pelatihan dan pengujian (Rios et al., 2021). Sampel data pelatihan adalah input di mana dengan menggunakan algoritma pembelajaran fitur dipelajari dalam pelatihan. Dalam pengujian, mesin eksekusi digunakan oleh algoritma pembelajaran untuk membuat prediksi untuk data pengujian yang tidak diketahui. Data terklasifikasi diberikan sebagai keluaran oleh model pembelajaran untuk mendeteksi serangan baru (Eliyan & Di Pietro, 2021). Machine Learning sendiri dapat mengaktifkan pencegahan dan perilaku prediktif DDoS, sehingga teknik ini sangat direkomendasikan untuk dapat menganalisis lalu lintas jaringan secara real-time, mengklasifikasikan, menentukan data normal, dan memprediksi serangan DDoS dengan akurasi tinggi (Cvitić et al., 2021).

Algoritma pada ML dapat diterapkan pada kumpulan data dengan atau tanpa seleksi fitur (Gaur & Kumar, 2021). Algoritma pembelajaran yang diawasi membangun model matematika dari sekumpulan data yang berisi input dan output yang diinginkan (D'hooge et al., 2020). Data tersebut dikenal sebagai data pelatihan dan terdiri dari sekumpulan contoh pelatihan. Setiap contoh pelatihan memiliki satu atau lebih input dan output yang diinginkan (Gopi et al., 2021). Ini juga dikenal sebagai sinyal pengawasan. Algoritma unsupervised learning mengambil sekumpulan data yang hanya berisi input dan menemukan pola dalam data, seperti pengelompokan atau pengelompokan titik data (Deka et al., 2019). Oleh karena itu, algoritma ML ini belajar dari data uji yang belum diberi label, diklasifikasikan, atau dikategorikan. Algoritma unsupervised learning mengidentifikasi kesamaan dalam data dan bereaksi berdasarkan ada atau tidak adanya kesamaan tersebut di setiap bagian data baru (Kim, 2019). Seleksi fitur memiliki dampak penting dalam meningkatkan akurasi model. Proses ini merupakan fase pertama dan krusial untuk meningkatkan akurasi prediksi, tingkat deteksi, dan untuk mengurangi waktu eksekusi model DDoS (Nimbalkar & Kshirsagar, 2021)(Bindra & Sood, 2020). Tahapan ML ini diawali dengan pengumpulan dataset yang kemudian diekstraksi ke csv agar bisa dianalisis dalam bentuk angka. Tahapan selanjutnya ada rekayasa fitur dimana, seleksi fitur berperan penting dalam memprediksi apakah nantinya masuk kategori serangan atau tidak. Gambaran tahapan ML ini ditunjukkan pada gambar 2.

Gambar 2. Tahapan Machine Learning

III. HASIL DAN PEMBAHASAN

A. Dataset

Dataset merupakan salah satu komponen penting pada ML. Kumpulan Data yang masih asli dan belum terkontaminasi ini merupakan dataset yang bagus. Salah satu dataset yang berisi intrusi jaringan adalah UNSW-NB 15 dibuat menggunakan alat IXIA PerfectStorm dan riset dari Australian Centre for Cyber Security (ACCS) untuk mengekstrak hibrida aktivitas serangan modern dan kontemporer dari lalu lintas jaringan. Alat tcpdump digunakan untuk menangkap 100 GB lalu lintas jaringan mentah (file pcap). Setiap file pcap berisi 1000 MB untuk analisis paket yang mudah. Teknik Argus dan Bro-IDS dan dua belas prosedur dijalankan dalam implementasi paralel untuk menghasilkan 49 fitur dengan label kelas yang ditunjukkan

pada tabel 1. Kumpulan data ini berisi 2.540.044 catatan yang disimpan dalam empat file CSV. Selain itu, bagian dari kumpulan data ini dibagi menjadi kumpulan pelatihan dan kumpulan pengujian. Set pelatihan melibatkan 175.341 catatan, sedangkan set pengujian berisi 82.332 catatan dari semua jenis serangan yang terlibat dan catatan normal. KDD99 sangat umum di kalangan peneliti dalam penelitian DDoS. Sebuah survei oleh Referensi (Özgür & Erdem, 2017) menemukan bahwa 142 studi telah menggunakan dataset KDD99 dari 2010 hingga 2015. Dataset ini berisi 41 fitur yang berisi 494.021 dan 311.029 catatan dalam set pelatihan dan pengujian.

Tabel 1. Fitur pada Dataset UNSW-NB15 dan KDD99

No	UNSW-NB15	Tipe	KDD99	Tipe
1	srcip	Flow	duration	Basic
2	sport	Flow	protocol_type	Basic
3	dstip	Flow	service	Basic
4	dsport	Flow	flag	Basic
5	proto	Flow	src_bytes	Basic
6	state	Basic	dst_bytes	Basic
7	dur	Basic	land	Basic
8	sbytes	Basic	wrong_fragment	Basic
9	dbytes	Basic	urgent	Basic
10	sttl	Basic	hot	Content
11	dttl	Basic	num_failed_logins	Content
12	sloss	Basic	logged_in	Content
13	dloss	Basic	lnum_compromised	Content
14	service	Basic	lroot_shell	Content
15	sload	Basic	lsu_attempted	Content
16	dload	Basic	lnum_root	Content
17	spkts	Basic	lnum_file_creations	Content
18	dpkts	Basic	lnum_shells	Content
19	swin	Content	lnum_access_files	Content
20	dwin	Content	lnum_outbound_cmds	Content
21	stcpb	Content	is_host_login	Content
22	dtpcb	Content	is_guest_login	Time
23	smeansz	Content	count	Time
24	dmeansz	Content	srv_count	Time
25	trans_depth	Content	serror_rate	Time
26	res_bdy_len	Content	srv_error_rate	Time
27	sjit	Time	rerror_rate	Time
28	djit	Time	srv_rerror_rate	Time
29	stime	Time	same_srv_rate	Time
30	ltime	Time	diff_srv_rate	Time
31	sintpkt	Time	srv_diff_host_rate	Time
32	dintpkt	Time	dst_host_count	Host
33	tcprrt	Time	dst_host_srv_count	Host
34	synack	Time	dst_host_same_srv_rate	Host
35	ackdat	Time	dst_host_diff_srv_rate	Host
36	is_sm_ips_ports	Time	dst_host_same_src_port_rate	Host
37	ct_state_ttl	Add	dst_host_srv_diff_host_rate	Host
38	ct_flw_http_mthd	Add	dst_host_serror_rate	Host
39	is_ftp_login	Add	dst_host_srv_serror_rate	Host
40	ct_ftp_cmd	Add	dst_host_rerror_rate	Host
41	ct_srv_src	Add	dst_host_srv_rerror_rate	Host
42	ct_srv_dst	Add		
43	ct_dst_ltm	Add		
44	ct_src_ltm	Add		

45	ct_src_dport_ltm	Add		
46	ct_dst_sport_ltm	Add		
47	ct_dst_src_ltm	Add		
48	Attack_cat	Label		
49	Label	Label		

Sumber Tabel (Al-Daweri et al., 2020)

2. Seleksi Fitur

Seleksi fitur ini dapat berdampak signifikan pada Akurasi, Seleksi Fitur adalah sebuah cara yang digunakan untuk dapat mengoptimalkan kinerja dari *classifier* (Somantri & Khambali, 2017). Tingkat Deteksi, dan waktu eksekusi sistem ML DDoS yang ada (Tonkal et al., 2021). Dalam konteks ini, ini berfokus pada penilaian kinerja banyak eksperimen DDoS berdasarkan seleksi fitur dan algoritma ML (Maslan et al., 2020). Dengan mempelajari kumpulan data DDoS yang berbeda, algoritme, dan hasil terbaru dari beberapa proyek penelitian, kami meninjau banyak strategi seleksi fitur yang diterapkan untuk memprediksi serangan DDoS. Pada Tabel 2 menunjukkan beberapa penelitian yang menyatakan bahwa seleksi fitur menjadi salah satu faktor untuk meningkatkan tingkat akurasi selain dari dataset, algoritma dan lainnya.

Proses seleksi fitur didasarkan pada subset seleksi fitur yang paling konsisten, relevan, dan tidak berlebihan dari vektor fitur (Dong et al., 2021). Ini tidak hanya mengurangi waktu pelatihan dan penyesuaian model tetapi pada akhirnya membantu mencegah overfitting (Aamir & Zaidi, 2019b). Seleksi fitur, sebagai teknik pengurangan dimensi, bertujuan untuk memilih sebagian kecil fitur yang relevan dari fitur asli dengan menghilangkan fitur yang tidak relevan, berlebihan, atau terkandung (Aamir & Zaidi, 2019a). Seleksi fitur biasanya dapat menghasilkan kinerja pembelajaran yang lebih baik, akurasi pembelajaran yang lebih tinggi, biaya komputasi yang lebih rendah, dan kemampuan interpretasi model yang lebih baik (Miao & Niu, 2016). Artikel ini berfokus pada proses seleksi fitur dan memberikan gambaran yang komprehensif dan terstruktur tentang jenis seleksi fitur, metodologi, dan teknik dari perspektif data dan algoritma (Venkatesh & Anuradha, 2019). Mengurangi ukuran kumpulan data dapat dilakukan dengan salah satu dari dua cara: seleksi fitur atau pengurangan kumpulan sampel. Tujuan dari teknik seleksi fitur antara lain (Rudnicki W.R., Wrzesień M., 2015):

- Model penyederhanaan untuk interpretasi yang lebih mudah oleh peneliti/pengguna
- Waktu pelatihan lebih singkat
- Hindari dimensional
- Meningkatkan generalisasi dengan mengurangi overfitting
- Kurangi ukuran set data untuk mencapai analisis yang lebih efisien
- Mengadaptasi kumpulan data yang paling sesuai dengan metode analisis yang dipilih.

Tabel 2. Ringkasan Penelitian Dampak Seleksi Fitur

Referensi	Dataset	Jumlah Fitur	Algoritma yang digunakan	Hasil
Belouch (Belouch et al., 2018)	NSL-KDD and UNSW-NB15	9 & 9	HFWFS	99.93% and 98.95%
Kunhare (Kunhare & Tiwari, 2018)	KDD99	41	NB, Random forest, J48	92.75%, 99.95%, 99.95%
Zhiqiang (Zhiqiang et al., 2019)	UNSW-NB15	49	Feed-forward ANN	99%
Bagui (Bagui et al., 2019)	UNSW-NB15	13	NB & J48	92.97% & 95.06%
Al-Daweri (Al-Daweri et al., 2020)	KDD99 & UNSW-NB15	36 & 12	D-CFA	99.40% & 86.57%
Kasongo (Kasongo & Sun, 2020)	UNSW-NB15	19	XGBoost-based feature selection (SVM, kNN, LR, ANN, DT)	90.85%
Arif (Wirawan et al., 2020)	UNSW-NB15	7	ANN	97.76%
Aleesa (Aleesa et al., 2021)	UNSW-NB15	48	ANN	99.26%

Geeta Kocher (Kocher & Kumar, 2021)	UNSW-NB15	49	RF	99.57%
Andi (Maslan et al., 2020)	Eksperimen	4	NB, RF, NN, SVM, KNN	97.96%, 98.70%, 98.70%, 98.14%, 97.63%
Polat (Polat et al., 2020)	Eksperimen	12	KNN	97.7%
Somasundaram (Somasundaram & Meenakshi, 2020)	KDD Cup	12	Proposed (EFSM)	99.79%
Aamir (Aamir & Zaidi, 2019)	Eksperimen	22	KNN	93.53%

IV. SIMPULAN

IDS (Intrusion Detection System) merupakan mekanisme keamanan yang dirancang untuk mencegah kejahatan siber seperti DDoS ini. Proses deteksi intrusi yang efisien sangat dipertimbangkan saat merancang model salah satunya dengan seleksi fitur. Seleksi fitur sangat penting dari perspektif penerapan algoritma ML. Seleksi fitur ini dapat meningkatkan kinerja algoritma, mengurangi kompleksitas dan waktu eksekusi. Dengan kata lain, model ML menjadi efisien dengan mengimplementasikan seleksi fitur. Ini menghapus fitur yang tidak relevan dan berlebihan tanpa kehilangan banyak informasi.

V. UCAPAN TERIMA KASIH

Penelitian ini dapat terlaksana dengan bantuan Dana dari Daftar Isian Pelaksanaan Anggaran (DIPA) Politeknik Negeri Cilacap Nomor DIPA-023.18.2.677629/2021.

DAFTAR PUSTAKA

- Aamir, M., & Zaidi, S. M. A. (2019a). DDoS attack detection with feature engineering and machine learning : the framework and performance evaluation. *International Journal of Information Security*, 18, 761–785. <https://doi.org/10.1007/s10207-019-00434-1>
- Aamir, M., & Zaidi, S. M. A. (2019b). DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation. *International Journal of Information Security*, 18(6), 761–785. <https://doi.org/10.1007/s10207-019-00434-1>
- Al-Daweri, M. S., Zainol Ariffin, K. A., Abdullah, S., & Md. Senan, M. F. E. (2020). An Analysis of the KDD99 and UNSW-NB15 Datasets for the Intrusion Detection System. *Symmetry*, 12(10), 1666. <https://doi.org/10.3390/sym12101666>
- Aleesa, A. M., Younis, M., Mohammed, A. A., & Sahar, N. M. (2021). Deep-intrusion detection system with enhanced UNSW-NB15 dataset based on deep learning techniques. *Journal of Engineering Science and Technology*, 16(1), 711–727.
- Ambusaidi, M. A., He, X., Nanda, P., & Tan, Z. (2016). Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm. *IEEE Transactions on Computers*, 65(10), 2986–2998. <https://doi.org/10.1109/TC.2016.2519914>
- Bagui, S., Kalaimannan, E., Bagui, S., Nandi, D., & Pinto, A. (2019). Using machine learning techniques to identify rare cyber-attacks on the UNSW-NB15 dataset. *Security and Privacy*, 2(6), 1–13. <https://doi.org/10.1002/spy2.91>
- Bärli, E. M., Yazidi, A., Viedma, E. H., & Haugerud, H. (2021). DoS and DDoS mitigation using Variational Autoencoders. *Computer Networks*, 199(June), 108399. <https://doi.org/10.1016/j.comnet.2021.108399>
- Belouch, M., Elhadaj, S., & Idhammad, M. (2018). A hybrid filter-wrapper feature selection method for DDoS detection in cloud computing. *Intelligent Data Analysis*, 22(6), 1209–1226. <https://doi.org/10.3233/IDA-173624>

- Bindra, N., & Sood, M. (2020). Evaluating the impact of feature selection methods on the performance of the machine learning models in detecting DDoS attacks. *Romanian Journal of Information Science and Technology*, 23(3), 250–261.
- Cepheli, Ö., Büyükçorak, S., & Karabulut Kurt, G. (2016). Hybrid Intrusion Detection System for DDoS Attacks. *Journal of Electrical and Computer Engineering*, 2016. <https://doi.org/10.1155/2016/1075648>
- Chowdhury, F. Z. (2017). Economic Denial of Sustainability (EDoS) Mitigation Approaches in Cloud : Analysis and Open Challenges. *IEEE International Conference on Electrical Engineering and Computer Science (ICECOS) 2017 Economic*, 206–211.
- Cvitić, I., Peraković, D., Periša, M., & Botica, M. (2021). Novel approach for detection of IoT generated DDoS traffic. *Wireless Networks*, 27(3), 1573–1586. <https://doi.org/10.1007/s11276-019-02043-1>
- D'hooge, L., Wauters, T., Volckaert, B., & De Turck, F. (2020). Inter-dataset generalization strength of supervised machine learning methods for intrusion detection. *Journal of Information Security and Applications*, 54, 102564. <https://doi.org/10.1016/j.jisa.2020.102564>
- Deka, R. K., Bhattacharyya, D. K., & Kalita, J. K. (2019). Active learning to detect DDoS attack using ranked features. *Computer Communications*, 145(May), 203–222. <https://doi.org/10.1016/j.comcom.2019.06.010>
- Dong, H., Sun, J., & Sun, X. (2021). A Multi-Objective Multi-Label Feature Selection Algorithm Based on Shapley Value. *Entropy*, 23(8), 1094. <https://doi.org/10.3390/e23081094>
- Eliyan, L. F., & Di Pietro, R. (2021). DoS and DDoS attacks in Software Defined Networks: A survey of existing solutions and research challenges. *Future Generation Computer Systems*, 122, 149–171. <https://doi.org/10.1016/j.future.2021.03.011>
- Feroz Khan, A. B., & Anandharaj, G. (2020). A Multi-layer Security approach for DDoS detection in Internet of Things. *International Journal of Intelligent Unmanned Systems*, 9(3), 178–191. <https://doi.org/10.1108/IJIUS-06-2019-0029>
- Fuertes, W., Tunala, A., Moncayo, R., Meneses, F., & Toulkeridis, T. (2017). Software-based Platform for Education and Training of DDoS Attacks using Virtual Networks. *2017 International Conference on Software Security and Assurance (ICSSA)*, 94–99. <https://doi.org/10.1109/ICSSA.2017.19>
- Gaur, V., & Kumar, R. (2021). Analysis of Machine Learning Classifiers for Early Detection of DDoS Attacks on IoT Devices. *Arabian Journal for Science and Engineering*. <https://doi.org/10.1007/s13369-021-05947-3>
- Gopi, R., Sathiyamoorthi, V., Selvakumar, S., Manikandan, R., Chatterjee, P., Jhanjhi, N. Z., & Luhach, A. K. (2021). Enhanced method of ANN based model for detection of DDoS attacks on multimedia internet of things. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-021-10640-6>
- Haider, S., Akhunzada, A., Mustafa, I., Patel, T. B., Fernandez, A., Choo, K. R., & Iqbal, J. (2020). A Deep CNN Ensemble Framework for Efficient DDoS Attack Detection in Software Defined Networks. *IEEE Access*, 8, 53972–53983. <https://doi.org/10.1109/ACCESS.2020.2976908>
- Hajar, S., Ali, A., Ozawa, S., Ban, T., Nakazato, J., Shimamura, J., Tun, U., & Onn, H. (2016). A Neural Network Model for Detecting DDoS Attacks Using Darknet Traffic Features. *IEEE 2016 International Joint Conference on Neural Networks (IJCNN)*, November 2014, 2979–2985.
- Hildebrand, C. (2021). *Latest NETSCOUT Threat Intelligence Report Highlights Unprecedented DDoS Attack Activity*. <https://www.netscout.com/blog/latest-netscout-threat-intelligence-report-highlights>
- Jaafar, G. A., Abdullah, S. M., & Ismail, S. (2019). Review of Recent Detection Methods for HTTP DDoS Attack. *Journal of Computer Networks and Communications*, 2019(12), 1–10. <https://doi.org/10.1155/2019/1283472>
- Kasongo, S. M., & Sun, Y. (2020). Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset. *Journal of Big Data*, 7(1), 1–20. <https://doi.org/10.1186/s40537-020-00379-6>
- Kim, M. (2019). Supervised learning-based DDoS attacks detection: Tuning hyperparameters. *ETRI Journal*, 41(5), 560–573. <https://doi.org/10.4218/etrij.2019-0156>
- Kocher, G., & Kumar, G. (2021). Analysis of Machine Learning Algorithms with Feature Selection for Intrusion Detection using UNSW-NB15 Dataset. *International Journal of Network Security & Its Applications*, 13(1), 21–31. <https://doi.org/10.5121/ijnsa.2021.13102>
- Kunhare, N., & Tiwari, R. (2018). Study of the Attributes using Four Class Labels on KDD99 and NSL-KDD Datasets with Machine Learning Techniques. *2018 8th International Conference on Communication Systems and Network Technologies (CSNT)*, November 2018, 127–131. <https://doi.org/10.1109/CSNT.2018.8820244>
- Lima Filho, F. S. De, Silveira, F. A. F., de Medeiros Brito Junior, A., Vargas-Solar, G., & Silveira, L. F. (2019). Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning. *Security and Communication Networks*, 2019, 1–15. <https://doi.org/10.1155/2019/1574749>

- Machaka, P., & Bagula, A. (2016). Using Exponentially Weighted Moving Average Algorithm to Defend Against DDoS Attacks. *IEEE 2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics*.
- Mahjabin, T., Xiao, Y., Sun, G., & Jiang, W. (2017). A survey of distributed denial-of-service attack, prevention, and mitigation techniques. *International Journal of Distributed Sensor Networks*, 13(12), 155014771774146. <https://doi.org/10.1177/1550147717741463>
- Marta, I. K. K. A., Hartawan, I. N. B., & Satwika, I. K. S. (2020). Analisis Sistem Monitoring Keamanan Server Dengan Sms Alert Berbasis Snort. *INSERT: Information System and Emerging Technology Journal*, 1(1), 25. <https://doi.org/10.23887/insert.v1i1.25874>
- Marvi, M., Arfeen, A., & Uddin, R. (2020). A generalized machine learning-based model for the detection of DDoS attacks. *International Journal of Network Management*, October, 1–22. <https://doi.org/10.1002/nem.2152>
- Maslan, A., Mohamad, K. M. Bin, & Mohd Foozy, F. B. (2020). Feature selection for DDoS detection using classification machine learning techniques. *IAES International Journal of Artificial Intelligence*, 9(1), 137–145. <https://doi.org/10.11591/ijai.v9.i1.pp137-145>
- Merouane, M. (2017). An approach for detecting and preventing DDoS attacks in campus. *Automatic Control and Computer Sciences*, 51(1), 13–23. <https://doi.org/10.3103/S0146411616060043>
- Miao, J., & Niu, L. (2016). A Survey on Feature Selection. *Procedia Computer Science*, 91(Information Technology and Quantitative Management (ITQM 2016)), 919–926. <https://doi.org/10.1016/j.procs.2016.07.111>
- Min, E., Long, J., Liu, Q., Cui, J., & Chen, W. (2018). TR-IDS: Anomaly-Based Intrusion Detection through Text-Convolutional Neural Network and Random Forest. *Security and Communication Networks*, 2018, 1–9. <https://doi.org/10.1155/2018/4943509>
- Moustafa, N., & Slay, J. (2016). The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. *Information Security Journal: A Global Perspective*, 25(1–3), 18–31. <https://doi.org/10.1080/19393555.2015.1125974>
- Nimbalkar, P., & Kshirsagar, D. (2021). Feature selection for intrusion detection system in Internet-of-Things (IoT). *ICT Express*, 7(2), 177–181. <https://doi.org/10.1016/j.icte.2021.04.012>
- Novaes, M. P., Carvalho, L. F., Lloret, J., & Proença, M. L. (2021). Adversarial Deep Learning approach detection and defense against DDoS attacks in SDN environments. *Future Generation Computer Systems*, 125, 156–167. <https://doi.org/10.1016/j.future.2021.06.047>
- Özgür, A., & Erdem, H. (2017). The impact of using large training data set KDD99 on classification accuracy. In *PeerJ* (Vol. 5). <https://doi.org/10.7287/peerj.preprints.2838>
- Parmar, H., & Gosai, A. (2015). Analysis and Study of Network Security at Transport Layer. *International Journal of Computer Applications*, 121(13), 35–40. <https://doi.org/10.5120/21604-4716>
- Polat, H., Polat, O., & Cetin, A. (2020). Detecting DDoS Attacks in Software-Defined Networks Through Feature Selection Methods and Machine Learning Models. *Sustainability*, 12(3), 1035. <https://doi.org/10.3390/su12031035>
- Ramadhan, G., Kurniawan, Y., Kim, C., Syn, A. T. C. P., & Ddos, F. (2016). Design of TCP SYN Flood DDoS Attack Detection Using Artificial Immune Systems. *IEEE 6th International Conference on System Engineering and Technology*, 72–76.
- Renuga Devi, R., & Umamaheswari, N. (2020). Detection of DDoS Attack in Cloud Computing using an Artificial Intelligence Based Approaches. *Journal of Critical Reviews*, 7(18), 2855–2861. <https://doi.org/10.31838/jcr.07.18.295>
- Rios, V. de M., Inácio, P. R. M., Magoni, D., & Freire, M. M. (2021). Detection of reduction-of-quality DDoS attacks using Fuzzy Logic and machine learning algorithms. *Computer Networks*, 186(December 2020), 107792. <https://doi.org/10.1016/j.comnet.2020.107792>
- Rudnicki W.R., Wrzesień M., P. W. (2015). Feature Selection for Data and Pattern Recognition. In U. Stańczyk & L. C. Jain (Eds.), *Feature selection for data and pattern classification* (Vol. 584). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-45620-0>
- Somantri, O., & Khambali, M. (2017). Feature Selection Klasifikasi Kategori Cerita Pendek Menggunakan Naïve Bayes dan Algoritme Genetika. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi (JNTETI)*, 6(3), 301–306. <https://doi.org/10.22146/jnteti.v6i3.332>
- Somasundaram, A., & Meenakshi, V. S. (2020). Ensemble Feature Selection Method and Combining Classifier Approach for DDoS Detection in Cloud. *International Journal of Advanced Science and Technology*, 29(3), 14348–14364.

- Sonule, A., Kalla, M., Jain, A., & Chouhan, D. S. (2020). Unsw-Nb15 Dataset and Machine Learning Based Intrusion Detection Systems. *International Journal of Engineering and Advanced Technology (IJEAT)*, 9(3), 2638–2648. <https://doi.org/10.35940/ijeat.C5809.029320>
- Thakkar, A., & Lohiya, R. (2021). A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions. *Artificial Intelligence Review*. <https://doi.org/10.1007/s10462-021-10037-9>
- Thapngam, T., Yu, S., Zhou, W., & Makki, S. K. (2014). Distributed Denial of Service (DDoS) detection by traffic pattern analysis. *Peer-to-Peer Networking and Applications*, 7(4), 346–358. <https://doi.org/10.1007/s12083-012-0173-3>
- Tonkal, Ö., Polat, H., Başaran, E., Cömert, Z., & Kocaoğlu, R. (2021). Machine Learning Approach Equipped with Neighbourhood Component Analysis for DDoS Attack Detection in Software-Defined Networking. *Electronics*, 10(11), 1227. <https://doi.org/10.3390/electronics10111227>
- Torabi, M., Udzir, N. I., Abdullah, M. T., & Yaakob, R. (2021). A Review on Feature Selection and Ensemble Techniques for Intrusion Detection System. *International Journal of Advanced Computer Science and Applications*, 12(5), 538–553. <https://doi.org/10.14569/IJACSA.2021.0120566>
- Venkatesh, B., & Anuradha, J. (2019). A Review of Feature Selection and Its Methods. *Cybernetics and Information Technologies*, 19(1), 3–26. <https://doi.org/10.2478/cait-2019-0001>
- Wirawan, A., Feresca, C., Foozy, M., & Azhari, A. (2020). Machine Learning-Based Distributed Denial of Service Attack Detection on Intrusion Detection System Regarding to Feature Selection. *International Journal Of Artificial Intelligence Research*, 4(1), 1–8. <https://doi.org/10.29099/ijair.v4i1.156>
- Zhang, J., Liu, P., He, J., & Zhang, Y. (2016). A Hadoop based analysis and detection model for IP Spoofing typed DDoS attack. *2016 IEEE TrustCom-BigDataSE-ISPA*, 1978–1985. <https://doi.org/10.1109/TrustCom.2016.300>
- Zhiqiang, L., Bing, L., & Zhijun, L. (2019). Modeling Network Intrusion Detection System Using Feed-Forward Neural Network Using UNSW-NB15 Dataset. *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)*, 299–303.
- Zhou, L., Liao, M., Yuan, C., & Zhang, H. (2017). Low-Rate DDoS Attack Detection Using Expectation of Packet Size. *Security and Communication Networks*, 2017(5), 1–14. <https://doi.org/10.1155/2017/3691629>