Intrusion Detection

Muhammad Ali Imron
Faculty of Information Technology
Institut Teknologi Batam
Batam, Indonesia
1822007@student.iteba.ac.id

***Abstract*—Dengan meningkatnya penggunaan sumber daya Internet, penyerang dunia maya menggunakan cara baru untuk menyerang layanan jaringan. Dengan demikian keamanan jaringan menjadi bagian tak terelakkan dari sistem jaringan. Untuk mendeteksi serangan tersebut secara efisien dan efektif, IDS yang kuat (Sistem Deteksi Intrusi) diperlukan. IDS adalah alat yang menganalisis setiap paket secara mendalam untuk mendeteksi aktivitas jahat dengan memantau jaringan atau sistem.Tujuan utama IDS adalah untuk mengidentifikasi atau tindakan abnormal dan untuk menginformasikan jaringan administrator tentang tindakan tersebut. Jadi, IDS adalah alat penting bagi administrator jaringan untuk mencegah jaringan dari serangan yang dikenal dan tidak dikenal yang membuat sumber daya jaringan lebih rentan.**

Metode pembelajaran mesin dapat digunakan untuk mempekerjakan sistem deteksi intrusi (IDS) yang efisien. Di dalam penelitian bekerja empat metode pembelajaran mesin adalah yang digunakan yaitu RF (Random Forest), DT (Decision Tree), MLP (Multilayer perceptron) dan SVM (Dukungan Vector Machine) untuk klasifikasi data. Dataset NSL KDD digunakan untuk melatih dan menguji ini berbagai model pembelajaran mesin. Pilihan fitur digunakan untuk menghilangkan yang tidak relevan dan tidak diinginkan fitur dari kumpulan data. Oleh karena itu pemilihan fitur mengurangi dimensi dataset yang pada gilirannya mengurangi kompleksitas komputasi. yang diusulkan keluaran model dievaluasi menggunakan tiga fitur subset, dipilih secara acak dari dataset NSL-KDD. Metode yang diusulkan memiliki akurasi klasifikasi lebih dari 99 persen.

***Keywords*—IDS ; RF ; DT ; SVM ; MLP ; Feature Selection; Machine learning.**

I. Pendahuluan

Serangan siber adalah tindakan jahat yang ditujukan untuk menargetkan jaringan dan sumber dayanya, untuk menghancurkan, menonaktifkan, mengubah, atau mendapatkan akses tidak sah ke sumber daya jaringan atau data yang mereka miliki [1] [1]. Kenaikan Serangan siber telah meningkatkan ancaman terhadap jaringan sumber daya yang mengarah pada tantangan baru bagi keamanan siber. Dengan munculnya teknologi baru seperti Internet hal-hal, komputasi cloud dan big data, the organisasi lebih rentan terhadap serangan ini. Oleh karena itu, sangat mendesak bagi organisasi untuk mengambil langkah-langkah yang diperlukan untuk melindungi pernyataan mereka dari kerusakan [2] [2]. Ini sangat penting untuk jaringan mengelola untuk mengambil mekanisme keamanan yang diperlukan untuk melindungi data dan sumber daya sensitif dari upaya yang tidak sah. Tujuan utama jaringan keamanan adalah untuk melindungi jaringan dari yang tidak diinginkan kode yang mengubah data, logika atau kode komputer yang dapat membahayakan sumber daya jaringan dan untuk menjaga integritas, meningkatkan ketersediaan dan melindungi kerahasiaan jaringan. Ada jenis serangan tertentu yang konvensional mekanisme keamanan tidak dapat mendeteksi karena penyusup menggunakan beberapa teknik yang berbeda, pendekatan untuk melewati dan menembus sistem keamanan jaringan. Firewall dan teknik Enkripsi tidak dapat memberikan solusi keamanan lengkap untuk semua jenis serangan jaringan (misalnya DoS). Secara otomatis firewall tidak dapat mempertahankan jaringan terhadap yang baru atau serangan yang tidak diketahui yang mungkin timbul. Firewall tidak bisa menjaga jaringan dari serangan berbahaya dilakukan oleh orang dalam dan serangan ini adalah dianggap paling merusak [3] [3]. Internet adalah penyerang yang terus berubah menemukan yang baru kerentanan, cara untuk menyerang jaringan. Jadi mekanisme keamanan baru diperlukan untuk menangani semua jenis serangan dengan cara yang efisien untuk mempertahankan keamanan jaringan komputer. Jadi disana ada kebutuhan untuk menyebarkan perangkat baru yang disempurnakan yang dapat mendeteksi semua jenis serangan intrusi dengan maksimal akurasi [4] [4]. Dengan demikian IDS memainkan peran penting dalam mendeteksi serangan semacam itu. Jika mekanisme keamanan seperti itu tidak diimplementasikan dalam jaringan penyerang attacker akan menyalahgunakan atau menghancurkan seluruh jaringan. Deteksi penyusupan adalah proses menemukan hal-hal yang mencurigakan pola dari data jaringan yang dapat merusak infrastruktur jaringan[5] [5]. Saat melakukan intrusi deteksi itu dibangun di atas fakta bahwa yang jahat lalu lintas terlihat berbeda dari lalu lintas normal. IDS dianggap sebagai mekanisme garis keamanan kedua, dirancang khusus untuk memeriksa semua lalu lintas jaringan dan sistem komputer, indra masuk dan keluar lalu lintas terus menerus untuk menemukan anomali tersembunyi dalam data dan segera menghasilkan peringatan jika sesuatu yang tidak biasa ditemukan di lalu lintas sebelumnya penyusup dapat merusak infrastruktur jaringan[6] [6]. Oleh karena itu fungsi utama dari deteksi intrusi adalah untuk memeriksa lalu lintas jaringan (baik yang masuk maupun yang keluar) dan mengambil tindakan yang tepat saat lalu lintas berbahaya diidentifikasi/ditemukan yaitu meningkatkan peringatan atau tindakan lain juga dimungkinkan seperti menjatuhkan paket. Komponen utama IDS adalah mesin pendeteksi dan fungsi utamanya adalah untuk menemukan lalu lintas berbahaya. Ketika beberapa sampel lalu lintas berbahaya telah diidentifikasi, IDS mengunci sampel ini di komponen lain yang disebut

log untuk digunakan nanti dan memutuskan tindakan yang mungkin dilakukan terhadap serangan yang dipilih untuk menjaga jaringan. IDS bisa dideskripsikan berdasarkan beberapa karakteristik yang akan menentukan taksonomi. Berdasarkan metode deteksi yang digunakan oleh mesin pendeteksi IDS dapat dibagi menjadi dua kategori [7] [7]: Anomali berbasis dan berbasis Penyalahgunaan. Kategori ini menentukan fungsi internal IDS. Dalam MIDS (atau tanda tangan atau berbasis pengetahuan) deteksi didasarkan pada realitas model yaitu model kami menjelaskan bagaimana serangan dan lalu lintas normal terlihat. AIDS atau berdasarkan perilaku model internal mendefinisikan struktur lalu lintas normal dan serang. Lebih lanjut berdasarkan penyebaran atau posisi IDS di arsitektur jaringan ini diklasifikasikan sebagai Network based (NIDS) melindungi seluruh jaringan yang diimplementasikan atau berbasis Host (HIDS) melindungi tuan rumah tunggal. Sebelum menyebarkan IDS di dunia nyata kinerja IDS harus dievaluasi. Jadi untuk tujuan evaluasi, peneliti membutuhkan kualitas dataset untuk melatih dan menguji model. Untuk mengevaluasi kinerja pembelajaran mesin pengklasifikasi pada kumpulan data yang diberikan metrik yang digunakan untuk tujuan evaluasi adalah akurasi, recall, presisi dll. Beberapa kumpulan data yang diketahui tersedia untuk umum adalah DARPA, KDD-CUP'99, NSL-KDD dan ADFA LD. Algoritma pembelajaran mesin telah digunakan berhasil di banyak bidang seperti Pemrosesan gambar, pemrosesan bahasa alami dan visi komputer dll. Algoritma pembelajaran mesin sangat bergantung pada data besar untuk menemukan pola tersembunyi dengan menggunakan set prosedur, fungsi transformasi kompleks[8] [8][9] [3]. Algoritma pembelajaran mesin menggunakan dua pembelajaran pendekatan untuk memahami data dengan jelas: Diawasi pembelajaran dan metode pembelajaran tanpa pengawasan. Di pembelajaran terawasi data yang digunakan untuk mempelajari model berisi data berlabel (data dengan output) tetapi dalam data pelatihan pembelajaran tanpa pengawasan tidak mengandung label model itu sendiri bersembunyi di dalam data untuk menemukan find pola alami. Selama fase pelatihan, data pelatihan digunakan untuk mengatur parameter dari fungsi yang kompleks, sehingga model dapat mengklasifikasikan data secara efisien. Penyusup mengubah perilaku mereka dengan menggunakan teknik dan alat terbaru. Penyusup menggunakan teknik seperti itu untuk mengubah perilaku jaringan mereka pola untuk melewati deteksi intrusi tradisional sistem. Sehingga menjadi perlu untuk penelitian komunitas untuk beralih ke yang baru dan dinamis pendekatan untuk mendeteksi dan mencegah intrusi ini. Oleh karena itu Menerapkan IDS yang efektif yang dapat mendeteksi serangan baru seperti itu adalah tugas yang menantang. Itu peningkatan cepat dalam teknik pembelajaran mesin telah meningkatkan prediksi dan daya komputasi dari mesin. Dengan demikian teknik ini dapat digunakan untuk membangun Sistem Deteksi Intrusi yang kuat. Baru saja peneliti menggunakan Sistem Deteksi Intrusi dan pemilihan fitur berdasarkan alat learning machine yang menunjukkan hasil yang menjanjikan dalam mendeteksi intrusi seperti Hutan Acak[10] [9], Pohon Keputusan[11] [10], Multilayer Perceptron[12] [11] dan vektor Dukungan Mesin[13] [12].

II. Literature Survey

Soodeh et al[14] [13] memperkenalkan pembelajaran mesin baru algoritma terdiri dari Algoritma Genetika, Logistik Regresi dan JST untuk deteksi intrusi sistem. Pada tahap pertama Regresi Logistik dan Algoritma genetika digunakan untuk mengekstrak yang berkorelasi subset fitur dari set data. Kemudian di detik Tahap Jaringan Syaraf Tiruan dilatih menggunakan PSO dan algoritma GS untuk mendeteksi intrusi dan. Dua kumpulan data digunakan untuk menilai kinerja model yang diusulkan yaitu NSL-KDD dan KDD piala'99. Model yang diusulkan mendapatkan tingkat akurasi yang lebih rendah tetapi model mendeteksi serangan lebih cepat dari yang lain metode berbasis ANN. Faezah et al [15] [14] mengurangi fitur data dengan menggunakan metode pembungkus berdasarkan Diferensial teknik evolusi untuk IDS. Jumlah fitur telah dikurangi karena fitur yang tidak relevan mempengaruhi keakuratan IDS. Idenya adalah untuk memilih beberapa fitur dari dataset NSL-KDD menggunakan evolusi diferensial dan menggunakan ETM untuk menentukan kinerja model yang diberikan. yang diusulkan model mencapai tingkat klasifikasi 80,15 persen untuk lima kelas dan 87,3 persen untuk dua kelas. Iram et al [16] [15] penelitian empiris tentang mesin pengklasifikasi pembelajaran berdasarkan SVM, KNN, LR, NB, MLP, RF, DT dan DLL untuk klasifikasi data jaringan sebagai abnormal dan normal digunakan dan kinerja penelitian dievaluasi pada empat himpunan bagian berbeda yang diturunkan dari NSL-KDD Himpunan data. Sebelum pelatihan model pelatihan data telah diproses sebelumnya berdasarkan fitur yang signifikan. Hasilnya mengungkapkan bahwa pembelajaran mesin pengklasifikasi menghasilkan hasil yang lebih baik untuk Denial of Service serangan dan hasil rendah dicapai untuk serangan U2R dan secara umum akurasi model adalah 99 persen. Miranal et al[17] [16] merancang IDS berdasarkan deep teknik pembelajaran menggunakan dataset NSL-KDD untuk mendeteksi intrusi di dalam jaringan. Model belajar sebagai serta memiliki kemampuan adaptif untuk menemukan pola baru yang tidak ditafsirkan sebelumnya. Model yang diusulkan menggunakan dataset NSL-KDD untuk pelatihan dan menggabungkan auto-encoder bersama dengan Regresi Logistik. Itu skor presisi yang dicapai oleh model lebih dari 84 persen. Yuyang et al [18] [17] mengusulkan IDS yang efisien berdasarkan pemilihan fitur heuristik yang disebut CFS-BA untuk mengurangi dimensi data berdasarkan korelasi antara atribut. Kemudian untuk tujuan deteksi pendekatan ensemble terdiri dari Random Forest, Forest dengan menghukum algoritma Atribut dan C4.5 dipekerjakan. Akhirnya dengan proses voting distribusi probabilitas peserta didik dasar adalah digabungkan untuk mengenali serangan. Hasil menunjukkan akurasi 99,8 persen dengan subset memiliki 10 fitur yang dipilih dari dataset NSL-KDD.

III. Metodologi

Bagian ini membahas hasil usulan kerja, di mana empat pengklasifikasi, RF, DT, SVM, dan MLP, digunakan untuk

menandai paket sebagai normal atau berbahaya berdasarkan data yang dikandungnya. Itu keluaran model dievaluasi menggunakan tiga subset fitur yang diambil dari set data NSL-KDD. Langkah-langkah yang digunakan dalam pekerjaan ini akan dijelaskan dan dirangkum dalam bagian berikut. NSL-KDD kumpulan data diproses sebelumnya pada fase pertama untuk dioptimalkan dan hapus fitur yang tidak perlu dari data mentah. Di tahap kedua, tiga set data dipilih di acak untuk menguji akurasi model. Mesin pengklasifikasi pembelajaran digunakan untuk pelatihan dan pengujian pada fase ketiga. Efek dari empat pengklasifikasi dievaluasi dalam tahap akhir.

A. Dataset and Preprocessing

Algoritma pembelajaran mesin bergantung pada jumlah besar data untuk melatih model sebelum mereka dapat menyediakan hasil yang lebih baik. Data biasanya disimpan dalam penyimpanan perangkat seperti file, database dll data ini tidak bisa langsung digunakan untuk tujuan pelatihan. Kita harus praproses atau perbaiki data untuk mencapai hasil yang lebih baik sebelum dapat diteruskan ke model pembelajaran mesin untuk tujuan pelatihan. Data pelatihan memungkinkan pengklasifikasi pembelajaran mesin untuk memahami bagaimana diberikan nilai berhubungan dengan kelas. Jadi mesin itu model pembelajaran dapat dengan mudah memahami data pelatihan dan memberikan hasil yang lebih baik. Langkah prapemrosesan data terdiri dari beberapa proses. Itu dimulai dari memuat data ke algoritma pembelajaran mesin untuk penanganan variabel dataset yang hilang, menskalakan data dengan bantuan standardisasi dan normalisasi, membagi dataset menjadi training dan testing dataset. Agar kita bisa meneruskan set pelatihan ke pengklasifikasi pembelajaran untuk tujuan pelatihan dan menggunakan set tes untuk mengevaluasi kinerja pengklasifikasi pembelajaran mesin. Tabel 1 merangkum detail dari tiga yang dipilih secara acak subset fitur dari set data NSL-KDD.

SELECTED FEATURE SUBSET	TOTAL NO. OF ROWS	NO. OF ROWS IN TRAINING SET	NO. OF ROWS IN TEST SET	SELECTED FEATURES	NO.OF FEATURES SELECTED
FEATURE SET 1ST	*1,25,971*	*81,882*	*44,089*	*7-9, 12, 20-24, 26, 28-34, 36-41.*	*23*
FEATURE SET 2ND	*1,25,971*	*81,882*	*44,089*	*23, 24, 26, 28-34, 37-41.*	*15*
FEATURE SET 3RD	*1,25,971*	*88,180*	*37,791*	*23, 24, 29-31, 34, 36-41.*	*12*

Gambar 1: Menunjukkan subset fitur yang dipilih dan jumlah instance yang dipilih secara acak dari Dataset NSL-KDD untuk pelatihan pengujian

B. Classification

Tantangan utama yang dihadapi IDS adalah false tingkat alarm (negatif palsu dan positif palsu) dan kekurangan dari respon waktu nyata. Algoritma pembelajaran mesin memiliki kekuatan untuk mengatasi tantangan seperti itu. Mesin teknik pembelajaran dapat digunakan untuk membangun kecerdasan IDS yang dapat mendeteksi baik yang diketahui maupun yang tidak diketahui serangan dengan kecepatan tinggi, akurasi maksimum dan tingkat alarm palsu minimum [19] [6] [20] [18]. Jadi mesin algoritma pembelajaran dapat digunakan untuk meningkatkan IDS dengan kemampuan yang ditingkatkan. pengklasifikasi pembelajaran mesin, yaitu RF (Random Forest), DT (Decision Tree), MLP (Multi-layer Perceptron), dan SVM (Dukungan Vector Machine) digunakan sebagai pemantau intrusi mesin untuk menemukan intrusi dengan mengklasifikasikan data. Hutan acak adalah pembelajaran terawasi yang kuat algoritma. Hutan acak adalah pengklasifikasi ansambel (terdiri dari multiple decision tress) yang digunakan untuk meningkatkan kinerja sistem[21] [19]. Itu output dari beberapa pohon dipilih untuk mengkategorikan data di kelas yang sesuai. Hutan Acak paling banyak menggunakan algoritma pembelajaran mesin yang diawasi untuk pengelompokan, klasifikasi data berdasarkan kesamaan fitur yang mereka bagikan. Hutan Acak dibangun dari pohon yang terbatas. Setiap pohon berperilaku seperti satu keputusan pohon di mana setiap pohon memilih fitur secara acak dari Himpunan data. Oleh karena itu, Menggunakan Pohon Acak untuk tujuan klasifikasi, jumlah pohon harus diperbaiki sebelum implementasi. DT adalah pembelajaran mesin algoritma klasifikasi untuk mengklasifikasikan data. Itu algoritma pohon keputusan secara predikatif dipelajari untuk membangun model dari kumpulan data yang dikategorikan ke memetakan sebuah instance berdasarkan serangkaian fitur yang dipilih ke kelas tertentu[22] [20]. Setiap sampel ditentukan oleh nilai-nilai mereka dari fitur masing-masing. Fungsi utama adalah untuk mendeteksi fitur, yang paling baik membagi data ke dalam kelasnya masing-masing. Node dapat dibagi dengan menggunakan entropi. Entropi mengukur kemurnian pemisahan sampel dalam simpul. Untuk tujuan membagi entropi digunakan dalam penelitian ini untuk memilih node terbaik. MLP (Multilayer Perceptron) adalah jaringan saraf terdiri dari satu atau lebih dari satu lapisan tersembunyi. Lapisan dalam MLP harus minimal tiga lapisan yaitu input, output dan lapisan tersembunyi yang memetakan variabel masukan menjadi keluaran akhir[5] [5]. Diperbaiki fungsi unit linier digunakan untuk melatih dan menguji model dan 10 neuron digunakan hanya di hidden lapisan. SVM adalah algoritma pembelajaran terawasi yang digunakan untuk kategorisasi biner dari non linier dan linier data. SVM (Support Vector Machine) mengklasifikasikan: data dengan membangun hyperplane N-dimensi dengan membagi sekelompok sampel positif dari sekelompok sampel negatif dengan margin tertinggi [4] [4][5] [5].Untuk pelatihan fungsi kernel basis radial diterapkan untuk memaksimalkan hasil prediksi untuk data non-linear. Kinerja model yang diusulkan adalah dianalisis pada tiga subset fitur berbeda yang diekstraksi dari kumpulan data NSL-KDD. Pra-pemrosesan adalah tugas penting untuk menghapus/mengganti fitur yang tidak relevan untuk meningkatkan akurasi dan meningkatkan ketahanan dari proses deteksi. Penampilan, biaya komputasi IDS didasarkan pada fitur/dimensi yang dipilih dari dataset, oleh karena itu, kumpulan data telah diproses sebelumnya untuk dihapus atribut yang tidak penting. Dalam penelitian ini bekerja dua dataset digunakan. Fitur-fitur yang berkontribusi yang paling untuk klasifikasi dipilih secara acak.

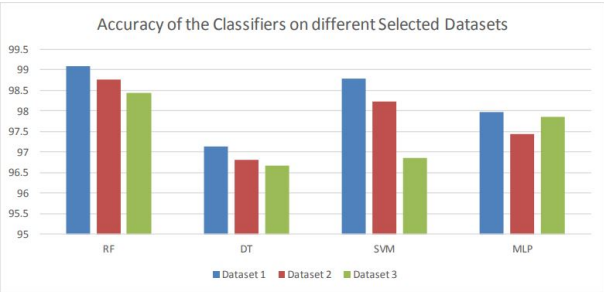
IV. Results

Percobaan dilakukan pada dataset NSL-KDD. Dataset NSL-KDD memiliki 41 kolom, membuatnya sulit untuk bek-

erja karena mereka meningkatkan biaya komputasi. Oleh karena itu, dataset direduksi untuk memenuhi persyaratan untuk percobaan. Tiga set data adalah dipilih secara acak dari dataset asli untuk mengurangi biaya komputasi. Semua percobaan adalah dilakukan di Google Colab dan efektivitas semua pengklasifikasi dalam mengklasifikasikan kumpulan data NSL-KDD adalah dipelajari. Pengklasifikasi RF menghasilkan yang tertinggi akurasi di atas 99 persen menggunakan subset fitur pertama. Random Forest adalah teknik klasifikasi ensemble (beberapa pengklasifikasi digabungkan untuk meningkatkan prediksi) yang menggunakan bilangan berhingga (dua atau lebih) pohon keputusan untuk melakukan klasifikasi. Jadi model mengurangi biaya komputasi dengan menghapus beberapa fitur yang tidak relevan dan meningkatkan akurasi model terutama untuk RF dan DT. Tabel 2 menunjukkan hasil pengklasifikasi yang berbeda di mengklasifikasikan data menggunakan tiga yang dipilih secara acak subset fitur dan gambar 1 menunjukkan grafik representasi temuan dalam hal akurasi pada berbagai subset fitur.

SR.NO.	CLASSIFIER	ACCURACY OF CLASSIFERS ON DIFFERENT SELECTED DATASETS / ATTRIBUTES		
		DATASET 1ST WITH 23 FEATURES	DATASET 2ND WITH 15 FEATURES	DATASET 3RD WITH 12 ATTRIBUTES
1.	RF	99.1%	98.77%	98.43%
2.	DT	97.14%	96.81%	96.66%
3.	SVM	98.79%	98.24%	96.85%
4.	MLP	97.97%	97.43%	97.85%

Gambar 2: Hasil yang dihasilkan oleh empat pengklasifikasi (RF, DT, SVM, MLP) pada tiga set data dengan set fitur yang berbeda



Gambar 3: Representasi grafis dari hasil yang dihasilkan

V. Kesimpulan

Dalam makalah ini, eksperimen empiris dilakukan menggunakan empat pengklasifikasi pembelajaran mesin yaitu RF, DT, MLP, dan SVM untuk menguji dan menguji efisiensi dan kinerja. Pelatihan dan penguji dilakukan pada tiga subset fitur yang diekstraksi dari set data deteksi intrusi NSL-KDD. Mulanya Dataset NSL-KDD telah dibawa sebelumnya untuk memilih fitur yang relevan untuk meningkatkan efisiensi dan mengurangi waktu pelatihan. Dalam Percobaan 81.882 kasus baris yang digunakan untuk melatih mesin yang dipilih model pembelajaran. Untuk tujuan pengujian 44.089 acak contoh yang digunakan. hasil berdasarkan yang dicapai hutan acak menghasilkan tingkat klasifikasi tertinggi lebih dari 99 persen, dan pohon keputusan menghasilkan tingkat akurasi terendah 96,60 persen di antara empat pengklasifikasian. Para peneliti harus fokus pada positif palsu dan kinerja negatif negatif yang menurunkan kinerja model deteksi intrusi. itu studi empiris telah mengungkapkan bahwa tidak ada pembelajaran mesin yang dapat mendeteksi semua bisa jenis serangan secara efektif. Di masa depan, relevan fitur dapat diekstraksi dari dataset asli untuk mengurangi waktu dan meningkatkan tingkat akurasi pengklasifikasi pembelajaran mesin. Metode berbasis ensemble dapat digunakan untuk menguji dan kinerja, metode ini dapat memprediksi serangan secara efisien.

References

[1] S. Bechhofer, M. T. Özsu, and L. Liu, "Owl: Web ontology language," in {*Encyclopedia of Database Systems*}. Springer Nature, 2009.

[2] F. Masoodi, S. Alam, and S. T. Siddiqui, "Security & privacy threats, attacks and countermeasures in internet of things," *International Journal of Network Security & Its Applications (IJNSA) Vol*, vol. 11, 2019.

[3] A. S. Ahanger, S. M. Khan, and F. Masoodi, "An effective intrusion detection system using supervised machine learning techniques," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 2021, pp. 1639–1644.

[4] C. Yang, G. N. Odvody, C. J. Fernandez, J. A. Landivar, R. R. Minzenmayer, and R. L. Nichols, "Evaluating unsupervised and supervised image classification methods for mapping cotton root rot," *Precision Agriculture*, vol. 16, no. 2, pp. 201–215, 2015.

[5] S. R. Sain, "The nature of statistical learning theory," 1996.

[6] F. S. Masoodi and M. U. Bokhari, "Symmetric algorithms i," in *Emerging Security Algorithms and Techniques*. Chapman and Hall/CRC, 2019, pp. 79–95.

[7] A. S. Ashoor and S. Gore, "Difference between intrusion detection system (ids) and intrusion prevention system (ips)," in *International Conference on Network Security and Applications*. Springer, 2011, pp. 497–501.

[8] H. Rajadurai and U. D. Gandhi, "A stacked ensemble learning model for intrusion detection in wireless network," *Neural computing and applications*, pp. 1–9, 2020.

[9] C. Ambikavathi, S. K. Srivatsa *et al.*, "Predictor selection and attack classification using random forest for intrusion detection," *Journal of Scientific and Industrial Research (JSIR)*, vol. 79, no. 05, pp. 365–368, 2020.

[10] I. H. Sarker, Y. B. Abushark, F. Alsolami, and A. I. Khan, "Intrudtree: a machine learning based cyber security intrusion detection model," *Symmetry*, vol. 12, no. 5, p. 754, 2020.

[11] M. Moukhafi, K. E. Yassini, and S. Bri, "Intelligent intrusion detection system using multilayer perceptron optimised by genetic algorithm," *International Journal of Computational Intelligence Studies*, vol. 9, no. 3, pp. 190–199, 2020.

[12] M. Safaldin, M. Otair, and L. Abualigah, "Improved binary gray wolf optimizer and svm for intrusion detection system in wireless sensor networks," *Journal of ambient intelligence and humanized computing*, vol. 12, no. 2, pp. 1559–1576, 2021.

[13] S. Hosseini, "A new machine learning method consisting of ga-lr and ann for attack detection," *Wireless Networks*, vol. 26, no. 6, pp. 4149–4162, 2020.

[14] F. H. Almasoudy, W. L. Al-Yaseen, and A. K. Idrees, "Differential evolution wrapper feature selection for intrusion detection system," *Procedia Computer Science*, vol. 167, pp. 1230–1239, 2020.

[15] I. Abrar, Z. Ayub, F. Masoodi, and A. M. Bamhdi, "A machine learning approach for intrusion detection system on nsl-kdd dataset," in *2020 International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE, 2020, pp. 919–924.

[16] S. Gurung, M. K. Ghose, and A. Subedi, "Deep learning approach on network intrusion detection system using nsl-kdd dataset," *International Journal of Computer Network and Information Security*, vol. 11, no. 3, pp. 8–14, 2019.

[17] Y. Zhou, G. Cheng, S. Jiang, and M. Dai, "Building an efficient intrusion detection system based on feature selection and ensemble classifier. june 2020," *Computer Networks*, vol. 174.

[18] M. Jabbar, R. Aluvalu *et al.*, "Rfaode: A novel ensemble intrusion detection system," *Procedia computer science*, vol. 115, pp. 226–234, 2017.

[19] L. Rutkowski, L. Pietruczuk, P. Duda, and M. Jaworski, "Decision trees for mining data streams based on the mcdiarmid's bound," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1272–1279, 2012.

[20] I. M. Alsmadi and A. AlEroud, "Sdn-based real-time ids/ips alerting system," in *Information Fusion for Cyber-Security Analytics*. Springer, 2017, pp. 297–306.