Detecting Cyber-Attacks with Intrusion Detection Systems Exploring Machine Learning Approaches

Priya Agrawal[1], Hemant Pal[2], Ritesh Joshi[3], Priyanka Jain[4]

[1]*Research Scholar, Department of Computer Science, Medi-Caps University, Indore, MP, India*
[2]*Assistant Professor, Department of Computer Science, Medi-Caps University, Indore, MP, India*
[3]*Assistant Professor, Department of Computer Applications, Medi-Caps University, Indore, MP, India*
[4]*Assistant Professor, Department of Computer Science, Medi-Caps University, Indore, MP, India*

KEYWORDS

Intrusion detection system, Machine learning.

ABSTRACT

increase in internet usage has been paralleled by a surge in cyber-attacks, many of which are novel and necessitate advanced detection mechanisms. Intrusion Detection Systems amuse oneself is critical bit part within keep track of network congestion which identify malicious activity. The detection of emerging threats requires the development of models using vast amounts of data to effectively distinguish between normal and anomalous traffic patterns. This has led to the growing appeal of machine learning algorithms that intensify predictive rightness of Intrusion Detection Systems. However, the high dimensionality of data poses significant challenges, particularly the "curse of dimensionality," which can degrade classification performance. This issue has prompted the adoption of feature selection and dimensionality reduction techniques to improve classification outcomes. In response to the increasing undivided attention about that application at supervised ML for Intrusion Detection Systems, this wrapper attending a sweeping look over of supervised learning algorithms along with their effectiveness in intrusion detection system. We review the core concepts of IDS, machine learning methodologies, and dimensionality reduction approaches. Additionally, we provide a detailed taxonomy that outlines the suitability of various algorithms for different IDS datasets, with a focus on the impact of feature selection on enhancing classification performance.

Introduction:

Cybersecurity exist a heterogeneous branch of knowledge dedicated to sheltering digital skeletons, tactful data together with passing on channels from spiteful pursuits, unaccredited access and systemic in perils. It skirts a gamut of dominions, embrace network matrix, cryptographic treaties, prevalence reciprocation, act toward intelligence, and security diminution.

Within an aeon at cyber threats, out of ultra-modern persistent threats onto ground zero exploits, cybersecurity amalgamates inventive mechanisms like as artifice machine learning together with blockchain onto embattle digital biodiversity.

In cybersecurity an Intrusion Detection System is apparatus which motif to keep track of and scrutinize network congestion conversely system ventures to pick out spiteful ventures or proposed action infringements. One time located the Intrusion Detection System give rise to vigilantes that notify structure controllers to those potential threatening remarks. Which is a censorious unit of contemporary network shielding strategies.

1. Types of Intrusion Detection System:

Type of Intrusion Detection System	Network based Intrusion Detection System	Host based Intrusion Detection System	Signature based Intrusion Detection System	Anomaly based Intrusion Detection System	Hybrid Intrusion Detection System	Protocol based Intrusion Detection System	Application Based Intrusion Detection System
Description	Scrutinize network congestion for unsure ventures or motifs.	Scrutinize activity on sole hosts or implement that cover file arrangement coverts and logs.	Pick outs threats utilize pre-judged motifs or signatures.	Utilizes machine learning or statistical replicas onto find deviations from stock deportment.	amalgamate trait at signature-based and anomaly-based noticing about all-inclusive protection.	Focal point onto protocol-specific traffic like as HTTP, FTP to locate breach of protocol degree.	Scrutinize appeal-level ventures for fishy behaviour.
Reinforce	Sheild total networks. Locate attacks across numerous structures.	Provides ingredient analysis for distinct structures. Locates insider threats.	Effectual for studied attacks. Quickly able to locate.	Locates tale or eighter zero-day attacks. Adapts to zestful territory.	Equilibrium strengths of many types. Reduced false positives.	Efficacious for safeguard web servers and specific act as assistance.	Safeguard censorious appeals like as databases. Locates misuse at the appeal level.
Fragility	May grapple together with encrypted congestion. High bandwidth usage can restrict staging.	Soaring resource consumption onto the host end. Lean to keep track of distinct device.	Cannot locate up to the minutes and undisclosed threats like as zero-day attacks.	soaring false positive rate. Essentials proper training of tutoring system.	Multiplex arrangements and preservation.	Finite scope to distinct protocols.	Essentials deep amalgamation together with applications. Eighter not detect network level attacks.

Using Machine Learning for Intrusion Detection Systems which look right onto an influential advance toward to intensify their aptness to locate both studied together zero-day threats. Unlike orthodox methods which rely onto fixed rules or signatures, Machine Learning based Intrusion Detection Systems grips motifs and data-driven replicas to vigorously pick out potentiality intrusions.

Immerse we in this far-reaching a look over paper on Machine Learning-based Intrusion Detection Systems that offering a subtle exploration of ultra-modern modes, evolving

provocations and time ahead superintendence's in Machine Learning leveraging for vigorous cybersecurity.

That research shows a look over of pragmatic twelve papers coming from directory of papers which assist to appreciate directed weighty heed in the direction of feature selection and dimensionality reduction strategies those mark to magnify classing production via mitigating the impact of extraneous traits.

Objective:

Leading of a look over this paper exist upon utterly analyse dynamic relationship between Intrusion detection system (IDS) and Machine learning. The focus is specifically on highlighting the significancy that arise from the combination of these two field for different cyber attacks revolutions we navigate the fast development in network security, personalized and identify the attack type and network care from attacks are jointly altering with traditional network security practices. This inquiry seeks to enhance the scholarly discourse on the integration of IDS and machine learning. It aims to provide valuable insights that can assist future research, IDS in navigating the challenging sector.

2.Evolution of intrusion detection system:

ERA	Key milestones	Methodologies	Gaps
1980's	predefined rules for detecting known attacks.	Statistical anomaly detection Rule-based systems	- Limited detection of unknown threats High dependency on predefined rules
1990's	exploring ML models for detecting anomalies and define threats.	Supervised learning like- neural networks, decision trees Unsupervised learning like- clustering	High false positives attack is difficult to detect
2000 to 2010's	Focus to improve accuracy. ML-based IDS became more prevalent with ensemble methods gaining traction.	Support Vector Machines (SVM) Bayesian networks, KDD Cup 1999 dataset Random Forest, Boosting Decision Trees	Limited scalability Feature extraction complexity Computational overhead for large datasets required for training
2010 to 2015 2015 to 2020	Increasing IDS with use of deep learning for better abstraction capabilities. Combination of traditional machine learning methods with signature-based.	Convolutional Neural Networks and Deep Neural Networks Generative models Recurrent Neural Networks	Deep models seen as black-box Training time and resource-intensive Balancing false positives and detection accuracy
2020 to present	AI-Driven & Real-Time Systems is better IDS systems.	Explainable AI Reinforcement Learning Federated Learning	In AI decisions Need for explainability Real-time detection challenges

3. Machine learning application in IDS sector: A Look Over

This a look over paper explores recent research publications to uncover the significant impact of IDS in the rapidly advancing field of Machine Learning. Table is included in this exploration, summarizing the finding obtained from each research publication This comprehensive exhibition not only showcase the significant progress achieved in the field of ids, but also strategically highlights the areas where further study is need. Our objective is to provide a comprehensive summary that will assist colleague researchers inside plot a route the compounded field belonging to machine learning together with its applications in cyber security using IDS. We hope that understanding of the subject serve as a source of inspiration for future study.

Machine learning applications: Review

Ref No	Title	year	Method	Conclusion	Research gap
[i]	A Survey on Machine Learning Techniques for Intrusion Detection Systems. [1]	2013	Neural Languages Support Vector Machines, Fuzzy Logic, Bayesian Networks, Decision Trees, Genetic Algo.	This study provides Machine learning to enable autonomous threat detection, reducing reliance on human analysts.	this paper highlight Hybrid systems are needed onto ameliorate noticing speed, rightness and bring down false alarms.
[ii]	Ensuring network security with a robust intrusion detection system using ensemble-based machine learning. [2]	2023	Forest technique outperforms existing methods with over 99% accuracy and improved evaluation metrics.	RF based ensemble mock up achieved superior accuracy and detection rates compared to other strategies to demonstrated by the evaluation of the proposed IDS.	Not many of canon can reliably detect each of types of ambush, and the proposed approach, tested on specific datasets, may have limited generalizability to other data sources.
[iii]	Anomaly detection for fault detection in wireless community networks using machine learning.[3]	2023	Arithmetical inspection they probe the potential of the dissimilar ML go toward to identify wanton entryway non-fulfilment that came in data troupe.	Traffic spikes outside the failure interval are likely noise, while VAE-based deep learning outperforms other methods but at a higher computational cost. Including CPU/memory features improves anomaly detection, with MAD offering simplicity and SHAP providing better explainability.	lacks discussion on the limitations of unsupervised ML approaches in wireless networks about fault detection and does not compare its anomaly detection techniques with existing methods in the field.

[iv]	Intrusion Detection System Using Machine Learning Algorithms.[4]	2022	Naïve Bayes, Support Vector Machine, K-Nearest Neighbour Datasets: UNSW-NB15, NSL-KDD	SVM algorithm proved the most effective in detecting various types of attacks. Time ahead duty will hub on amend the processing time together implementing the model in real-time firewalls.	Real-time implementation of SVM in a firewall is not tested. - Further optimization of the model is required, especially in terms of processing time.
[v]	Enhancement of Intrusion Detection System using Machine Learning.[5]	2023	Machine Learning (Logistic Regression, SVM, Naive Bayes, AdaBoost) - Ensemble Model: AdaBoost with Logistic Regression	Intrusion Detection Systems using machine learning replicas as like Logistic Regression, Naïve Bayes and SVM, with an ensemble model (AdaBoost with Logistic Regression) achieving high accuracy	- Need for testing the proposed ensemble model in real-time environments - Performance improvement on specific attack types such as U2R could be addressed further.
[vi]	Unveiling Machine Learning Strategies and Considerations in Intrusion Detection Systems: A Comprehensive Survey.[6]	2024	machine learning and deep learning methods used in Intrusion Detection Systems	This study demonstrates about effectiveness of hybrid models in intrusion detection, achieving remarkable accuracy rates.	more comprehensive datasets and further improvement in reducing false positives
[vii]	Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning.[7]	2023	Random Forest, AdaBoost, XGBoost, and LightGBM with ensemble learning (bagging and boosting) to detect and classify insider threats. A customized CERT dataset was processed through model training and testing phases for effective threat identification.	ensemble methods meaningful ameliorate the showing of insider threat noting compared to individual replicas. effective experimental results with higher accuracy in classifying insider attacks using the	need for well-ordered proceed toward to pick out atypical happening related onto privilege escalation in a unified dataset, which this study addressed. Lack of proper identification of insider attacks and often used different

				proposed ensemble algorithms.	models on separate datasets.
[viii]	A Unified Deep Learning Anomaly Detection and Classification Approach for Smart Grid Environments. [8]	2021	This paper presents MENSA, an IDS using Autoencoder-GAN architectonics onto locate oddities and relegate cyber pounce in Smart Grid environments. corroborated into four factual SG territory, MENSA outperforms other ML and DL methods in accuracy, FPR, TPR, and F1 score.	MENSA effectively addressed oddity's location and cyber pounce classing into well turned-out Grid territories, outperforming other ML together DL methods in key metrics. It successfully detected 14 Modbus/TCP and 5 DNP3 cyberattacks, demonstrating scalability across protocols.	While the MENSA model demonstrated high performance, the paper highlights the need for further exploration of other industrial control protocols like Profinet and Ether Cat. Additionally, the paper suggests further research into optimization techniques to mitigate cyberattacks and association rules to correlate DL model outcomes.
[ix]	Machine Learning in Network Anomaly Detection: A Survey.[9]	2021	- Surveys four ML models: Supervised, Unsupervised, Semi-supervised, and Reinforcement Learning. - Discusses application of ML in different network environments.	- ML is effective and flexible for anomaly detection across diverse networks. - Emphasizes the importance of ML in addressing complex and evolving network threats.	- Lack of real-world implementation and Computational efficiency issues such as CPU usage, training time are not well studied. so, Next-gen networks require further exploration, especially with deep learning.
[x]	Comparative Research on Network Intrusion Detection Methods Based on Machine Learning.[10]	2022	This paper compares multiple ML canons, embrace decision trees, Naive Bayes, random forests, XGBoost, CNNs, SVM, and RNNs. The experiments use KDD CUP99 and NSL-KDD datasets.	Ensemble learning algorithms generally performed better in intrusion detection, while Naive Bayes excelled at detecting novel attacks. Deep learning algorithms required further refinement.	The need for better preprocessing techniques for different datasets and further study of deep learning models to optimize hyperparameters and improve results.
[xi]	Machine Learning Security Attacks and Défense Approaches for	2022	That paper surveys profuse machine learning safe future pounces, including data bank and replica poisoning, privacy	The study dispenses a peril model for ML safe future into cyber physical systems, details different	The paper highlights challenges in securing ML models, especially in preprocessing for diverse datasets, and

	Emerging Cyber Physical Applications: A Comprehensive Survey.[11]		breaches, and runtime disruptions. It also reviews Défense mechanisms to safeguard ML models.	attack mechanisms, and discusses Défense strategies. It emphasizes the need for robust Défense's in ML systems.	suggests further research into advanced Défense strategies to handle emerging threats.
[xii]	Machine learning based smart intrusion and fault identification (SIFI) in inverter based cyber-physical microgrids.[12]	2023	Development of ML based SIFI replica Use of ensemble classifiers Implementation of the model in the RAP Sim simulator Testing against DoS and MDI cyberattacks Comparison with SVM, RF, and NB classifier.	proposed SIFI method is effective in classifying and localizing various cyber-physical anomalies, including physical faults and specific cyber threats such as DoS attacks and malicious data injections.	Vulnerability due to the data sharing process in cooperative control methods, Challenges in identifying anomalies within control structures, Limitations of conventional algorithms (SVM, ANN, DWT), Lack of consideration for physical and cyber abnormalities from DoS threats in existing literature.

5.challenges and ethical considerations:

The incorporation of machine learning in intrusion detection system (IDS) application raises intricate ethical concerns within the complex realm of the network security attack. protecting from attack presents a significance challenge requiring through way to tackle concerns around network security and preventions. Furthermore, as machine learning methods and algorithms become essential in decision making processes, significant ethical considerations arise, particularly in addressing in algorithmic bias and guaranteeing the interpretability of these model. It is crucial to find a middle ground between promoting new ideas and ensuring ethical accountability as we traverse the complex landscape of intrusion detection system.

6. Future direction and innovation:

In this article, we explore the promising field of developing technologies that are set revolutionize the future of machine learning .In the context of network security authors are exploring that latest break throughs in technology ,such as innovative uses of machine learning sophisticated methods and algorithms and improved techniques for network attacks in addition to their advanced technological capabilities ,we thoroughly analysed the possible effect of these advances on network security using intrusion detection system ,envisioning a future where accuracy ,efficiency , and customised inventions come together .These developing methods and algorithms have the potential to revolutionaries network security techniques , leading to improved secure network and a more advanced and focus on intrusion detection system with use of machine learning.

7.conclusion:

In conclusion this extensive analysis of the "Intrusion detection system for security ", emphasize the adaptability of machine learning within meadow at intrusion detection system,

escorted by a determined focus onto its use in network security attacks. ML techniques, Ensemble models like Random Forest, AdaBoost, and hybrid approaches, are shown to significantly improve accuracy and detection rates in intrusion detection systems. These techniques identifying security threats, reducing the reliance on human intervention and improving the overall performance of IDS in various network environments. This comprehensive analysis of Intrusion Detection Systems pinnacles the adaptability of Machine Learning crafts in intensifying network certainty. ML methods, particularly ensemble models like Random Forest and AdaBoost, significantly improve accuracy and detection rates by automating threat identification and reducing human involvement. Canons such as Decision Trees, Support Vector Machines and Naive Bayes are particularly effective for specific tasks, although real-time application remains challenging due to processing and efficiency limitations. These review research also underscores the potential of like Software-Defined Networks (SDN) and Smart Grids, though optimizing for new attack types remains a gap. Additionally, there is a need for more comprehensive datasets to improve model generalization across diverse network environments. Hybrid models, combining anomaly-based and signature-based approaches, are seen as essential for balancing detection accuracy and minimizing false positives, especially against zero-day attacks. This field requires more real-world testing, optimization for computational efficiency, and exploration of increased many as well as various data bank onto address current limitations within enhance system reliability.

References

[1] Singh, J., & Nene, M. J. (2013). A Survey on Machine Learning Techniques for Intrusion Detection Systems. *International Journal of Advanced Research in Computer and Communication Engineering*, 2(11), 4349-4355.

[2] Md. Alamgir Hossain, Md. Saiful Islam, "Ensuring network security with a robust intrusion detection system using ensemble-based machine learning," *Array*, vol. 19, 2023, p. 100306

[3] Cerdà-Alabern, L., Iuhasz, G., & Gemmi, G. (2023). Anomaly detection for fault detection in wireless community networks using machine learning. *Computer Communications*, 202, 191-203.

[4] Tahri, R., Balouki, Y., Jarrar, A., & Lasbahani, A. (2022). Intrusion Detection System Using Machine Learning Algorithms. *ITM Web of Conferences*, 46, 02003.

[5] Mukesh Kumar Yadav, & Mahaiyo Ningshen. (2023). *Enhancement of Intrusion Detection System using Machine Learning*. *International Journal of Engineering Research & Technology (IJERT)*, 12(1), 170-179.

[6] Ali AH, Charfeddine M, Ammar B, Hamed BB, Albalwy F, Alqarafi A, Hussain A (2024). Unveiling machine learning strategies and considerations in intrusion detection systems: a comprehensive survey. *Frontiers in Computer Science*:138.

[7] Mehmood, M., Amin, R., Muslam, M. M. A., Xie, J., & Aldabbas, H. (2023). Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning. *IEEE Access*, *11*, 46561-46575.

[8] Siniosoglou, I., Radoglou-Grammatikis, P., Efstathopoulos, G., Fouliras, P., & Sarigiannidis, P. (2021). A Unified Deep Learning Anomaly Detection and Classification Approach for Smart Grid Environments. *IEEE Transactions on Network and Service Management*, 18(2), 1137-1150.

[9] Wang, S., Balarezo, J. F., Kandeepan, S., Al-Hourani, A., Chavez, K. G., & Rubinstein, B. (2021). *Machine learning in network anomaly detection: A survey*. IEEE Access, 9, 152379-152393

[10] Zhang, C., Jia, D., Wang, L., Wang, W., Liu, F., & Yang, A. (2022). Comparative research on network intrusion detection methods based on machine learning. *Computers & Security*, 121, 102861.

[11] Singh, J., Wazid, M., Das, A. K., Chamola, V., & Guizani, M. (2022). Machine learning security attacks and defense approaches for emerging cyber physical applications: A comprehensive survey. *Computer Communications*, 192, 1-25.

[12] Divya, R., Umamaheswari, S., & Stonier, A. A. (2023). Machine learning based smart intrusion and fault identification (SIFI) in inverter based cyber-physical microgrids. *Expert Systems with Applications, 238*, 122291.