Global Research journal of Natural Science and Technology (Grjnst)
Volume: 03 - Issue 3 (2025)
p.ISSN: 2790-7643 e.ISSN: 2790-7651

Optimizing Cybersecurity Threat Detection Using Machine Learning: A Comparative Study of Supervised and Unsupervised Approaches

Areeba Naseem Khan
(Correspondence)
cydrahmuneer@gmail.com
COMSATS, Attock Campus, Pakistan

Muhammad Saad Sarfraz Khan
Electrical Engineering Department,
COMSATS University, Lahore Pakistan

Muhammad Nawaz Khan
Institute of Engineering Mathematics,
University Malaysia Perlis (UniMAP)
Malaysia

Laiba Khawaja
Department of Software Engineering, FICT
Balochistan University of Information Technology Engineering and Management Sciences,
Balochistan Pakistan

Abstract:

As cyber-attacks become more sophisticated, conventional rule-based security systems are no longer adequate for effective threat detection in a timely manner. In this study, the effectiveness of machine learning (ML) algorithms to detect and respond to cybersecurity threats is compared using supervised and unsupervised learning approaches. Models are trained on benchmarking datasets like CICIDS2017 and NSL-KDD to analyze detection rate, false-positive rates, and computational complexity. Results show that supervised models such as Random Forest and Support Vector Machine are more accurate compared to unsupervised models, but the clustering-based methods have strong zero-day attack detection anomaly. These results validate a hybrid model that incorporates the strengths of both learning paradigms for designing future cybersecurity frameworks.

1. Introduction

Cybersecurity is a rising global concern as the reliance on digital infrastructure increases in critical sectors such as finance, health care, defense, and public administration. The increased interconnectivity of devices, systems, and services due to factors such as cloud computing, Internet of Things (IoT), and teleworking has presented a broader attack surface for cyber attackers (Conti et al., 2018; Yang & Liu, 2021). Modern-day companies are continuously besieged by advanced persistent threats (APTs), polymorphic malware, phishing attacks, denial-of-service (DoS) attacks, ransomware, and zero-day exploits (Ali et al., 2020; Nguyen & Armitage, 2008). Threats not only compromise sensitive information and operations but also national security and public safety (Buczak & Guven, 2016).

Historical cybersecurity controls, particularly signature-based Intrusion Detection Systems (IDS), operate in accordance with known patterns and attack signatures. While effective against previously mentioned threats, such systems are deficient in identifying new, obfuscated, or adaptive attack models (Patcha & Park, 2007; Depren et al., 2005). Their reactive nature also limits their evolution in real-time against changing threats. Machine learning (ML) technologies introduce a paradigm shift through data-driven, proactive defensive measures that can learn behavior patterns, categorize anomalies, and respond to dynamic threats with little human interaction (Aburomman & Reaz, 2016; Ahmed et al., 2016).

The recent past has witnessed several ML algorithms being researched by the cybersecurity fraternity, from the conventional supervised algorithms like Decision Trees and Support Vector Machines (SVM) to unsupervised ones like K-Means clustering, Isolation Forests, and Autoencoders, for their capacity to enhance threat detection mechanisms (Mukkamala et al., 2002; Siddiqui & Naeem, 2021). The deployment of these algorithms in practical applications, however, is accompanied by trade-offs. While supervised models will show improved accuracy using known-labeled data, they do not have flexibility when determining unseen attacks (Zhang et al., 2008; Choudhary et al., 2020). In contrast, unsupervised models can identify anomalies and outliers without the use of labeled data, thus being suitable for zero-day attack detection but typically at the cost of greater false positives and greater computational expense (Ring & Wunderlich, 2020; Patil & Thorat, 2019).

In this context, the present study attempts to investigate the use of machine learning algorithms for detecting cybersecurity threats. More specifically, it tries to respond to two critical research questions:

(a) To what extent do different machine learning techniques recognize known as well as hitherto unknown (zero-day) cyber attacks on realistic benchmark datasets?

(b) How do supervised and unsupervised machine learning models compare in relative trade-offs among detection accuracy, false-positive rate, capability to generalize, and system overhead (e.g., training time and computational resource usage)?

By answering these questions, this research aims to generate a comprehensive evaluation of ML-based threat detection systems so that it can provide guidance on future development of hybrid or adaptive cybersecurity mechanisms that achieve the optimal balance between accuracy, efficiency, and responsiveness (Islam & Abawajy, 2020; Kim et al., 2014).

2. Literature Review

The recent developments in machine learning (ML) have significantly influenced the design and functions of Intrusion Detection System (IDS), updating them from traditional, rule-based systems (Buczak & Guven, 2020; Shaukat et al., 2020). Various studies have explored the integration of ML algorithms to facilitate the detection of signature-based and anomalous cyber attacks. Chandola et al. (2009) and Sommer & Paxson (2010) laid the foundation with the reporting of advantages of machine learning and data mining techniques for high-dimensional cybersecurity data anomaly detection. They focused on system activity modeling and detecting anomalies from normal behavior patterns (Patcha & Park, 2007; Li & Li, 2017).

Supervised learning algorithms like Random Forests, Support Vector Machines (SVM), Decision Trees, and Gradient Boosted Trees have been highly effective in the detection of recognized patterns of attacks. These algorithms function by learning pre-classified datasets consisting of benign and malicious traffic instances that are labeled, enabling the system to develop effective decision boundaries (Gaurav & Bhardwaj, 2017; Das & Sharma, 2018). Sahu and Panda (2020), for example, demonstrated that Random Forest classifiers could identify with over 95% precision on benchmark sets such as NSL-KDD far better than linear models in terms of simplicity (Dhanabal & Shantharajah, 2015; Choudhary et al., 2020). While these benefits, supervised approaches completely rely on the availability of well-

tagged, balanced, and fresh datasets, wishful thinking because of the rapidly evolving nature of cyber threats (Ring et al., 2019; Sharafaldin et al., 2018).

Conversely, unsupervised learning algorithms such as k-Means Clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), Isolation Forests, and Autoencoders have been found to be viable alternatives (Ring & Wunderlich, 2020; Latah & Toker, 2021). They don't require labeled training data and are therefore well-suited to detect zero-day or otherwise unknown attacks (Sundaram & Ramamurthy, 2020). Zhang et al. (2019) investigated Autoencoder-based anomaly detection and found they could learn complex network behaviors for outlier detection (Vidyarthi et al., 2019). However, unsupervised approaches yield high false positives, leading to alert saturation and reduced system reliability in deployment environments (Beghdad, 2020; Le & Bassett, 2018).

To overcome the limitation of single-paradigm models, hybrid IDS architectures have been proposed integrating both supervised and unsupervised learning paradigms. They attempt to utilize the advantages of both approaches—utilizing supervised learning for known attack classification and unsupervised techniques for zero-day or anomaly attack detection (Alazab et al., 2012; Wang et al., 2017). Sultana et al. (2019) proposed a two-phase IDS where identified threats are screened out with a supervised classifier and unsupervised clustering algorithms classify anomalous outliers, resulting in improved overall detection resilience (Gu et al., 2016; Vinayakumar et al., 2017). Nonetheless, most hybrid systems lack comprehensive benchmarking across existing, real-world data, and it is uncommon for research to conduct systematic performance experiments that take into account scalability, flexibility, and false alarm trade-offs under homogeneous conditions (Islam & Abawajy, 2020; Hodo et al., 2017).

Against this background, this current study aims to fill this significant gap by carrying out a comparative assessment of supervised, unsupervised, and hybrid ML models with recent datasets such as CICIDS2017 and NSL-KDD (Sharafaldin et al., 2018; Ring et al., 2019). By comparing the same models under equivalent conditions, this latest research aims to provide practical insight into the usability and efficacy of ML-based IDS in modern cybersecurity environments (Siddiqui & Naeem, 2021; Zhang et al., 2020).

3. Methodology

The research employs an empirical experimental approach seeking to assess and compare the performance of supervised and unsupervised machine learning models for intrusion detection in cybersecurity. The experiments were carried out on two of the most widely acknowledged benchmark datasets, namely NSL-KDD and CICIDS2017, both of which include a comprehensive variety of cyber-attacks such as Denial-of-Service (DoS), Distributed Denial-of-Service (DDoS), User-to-Root (U2R), Remote-to-Local (R2L), and probing attacks. These datasets have been extensively used in intrusion detection research due to their richness, labeled instances, and representativeness of real-world network traffic patterns (Dhanabal & Shantharajah, 2015; Sharafaldin et al., 2018; Ring et al., 2019). Unlike the outdated KDD'99 dataset, which has redundant records and skewed distributions, NSL-KDD offers a more balanced benchmark, whereas CICIDS2017 provides realistic, contemporary attack scenarios, including botnets and web attacks, making them suitable for evaluating modern IDS systems (Ring et al., 2019; Siddiqui & Naeem, 2021).

3.1 Data Preprocessing

Before the models were trained, extensive preprocessing was performed to ensure data quality and machine learnability:

Normalization: Min-max normalization was applied to scale continuous features into a standardized range [0, 1]. This step enhances model stability and facilitates faster convergence, especially for algorithms sensitive to feature scaling such as SVM and k-Means (Choudhary et al., 2020; Gharib et al., 2016).

Feature Selection: Information Gain (IG) was employed to select the most informative attributes. Reducing dimensionality helps eliminate redundant features and decreases computational costs while preserving the most predictive signals in the data (Gu et al., 2016; Zhang et al., 2020).

Imbalanced Classes Handling: Both datasets suffer from class imbalances (attack records vs. normal traffic). To mitigate this, the Synthetic Minority Oversampling Technique (SMOTE) was applied, increasing minority-class representation and reducing model bias toward majority classes. This step is crucial for improving Recall and F1-Scores in IDS research (Ali et al., 2020; Shaukat et al., 2020).

3.2 Model Implementation

The chosen machine learning algorithms were classified into supervised and unsupervised methodologies to facilitate a comparative analysis.

Supervised Learning Models:

Random Forest (RF): An ensemble-based classifier that leverages bagging and feature randomness to achieve high accuracy and robustness. It has been widely validated for intrusion detection due to its interpretability and resilience against overfitting (Zhang et al., 2008; Choudhary et al., 2020).

Support Vector Machine (SVM): A strong classifier that uses kernel functions to transform inputs into high-dimensional feature spaces. Though computationally expensive, it has proven effective in complex, non-linear classification tasks (Mukkamala et al., 2002; Buczak & Guven, 2016).

Logistic Regression (LR): A simple linear model employed here as a baseline due to its efficiency and interpretability, though less effective in modeling non-linear attack behaviors (Das & Sharma, 2018).

Unsupervised Learning Models:

k-Means Clustering: A partitioning-based clustering algorithm that groups data into k clusters on the basis of similarity. While simple and scalable, its performance is sensitive to cluster initialization (Ring & Wunderlich, 2020; Patil & Thorat, 2019).

Isolation Forest: A tree-based anomaly detection method that isolates outliers using random feature splits. Its efficiency in detecting rare intrusions has been reported in recent IDS studies (Latah & Toker, 2021).

Autoencoders: Neural networks designed for input reconstruction. They capture normal traffic distributions and highlight deviations as anomalies, making them effective in detecting zero-day attacks (Sundaram & Ramamurthy, 2020; Siddiqui & Naeem, 2021).

3.3 Evaluation Metrics and Validation

Each model was evaluated with a standardized set of performance measures: Accuracy, Precision, Recall, F1-Score, False Positive Rate (FPR), and Area Under the ROC Curve (AUC). These measures provide a holistic understanding of model performance across detection, reliability, and misclassification risk (Buczak & Guven, 2016; Latah & Toker, 2021). In addition, a 10-fold cross-validation approach was employed to ensure generalizability and reduce the risk of overfitting, in line with best practices in IDS research (Le & Bassett, 2018; Zhang et al., 2020).

4. Results and Discussion

Experimental findings uncover differentiated performance trends among supervised and unsupervised models when tested on the two datasets, highlighting the strengths and weaknesses of each approach.

4.1 Supervised Models

Among the supervised models, Random Forest performed best, achieving 97.3% accuracy on the CICIDS2017 dataset. It exhibited a low false-positive rate (2.1%), demonstrating its efficacy in distinguishing benign from malicious traffic. Its AUC score of 0.98 indicates a strong classification ability, consistent with earlier studies where RF outperformed linear classifiers in IDS contexts (Zhang et al., 2008; Choudhary et al., 2020). SVM achieved slightly lower accuracy (95.5%) but maintained robust detection properties. However, its computational expense during training and prediction was significantly higher, especially on large datasets, which echoes the computational concerns highlighted by Buczak and Guven (2016) and Mukkamala et al. (2002). Logistic Regression, though computationally efficient

and interpretable, lagged in recognizing sophisticated, non-linear attack patterns, confirming its limitations for complex intrusion detection (Das & Sharma, 2018).

4.2 Unsupervised Models

Unsupervised models showed relatively lower detection accuracy but were valuable in identifying anomalies and potential zero-day attacks. The Autoencoder performed best among them, with 88.4% accuracy and an AUC of 0.91, demonstrating its ability to reconstruct normal traffic patterns and flag deviations. This aligns with findings by Sundaram and Ramamurthy (2020) and Siddiqui and Naeem (2021), who reported the promise of deep learning autoencoders in anomaly-based IDS. k-Means Clustering achieved 83.9% accuracy, but with a high false-positive rate, reflecting its sensitivity to cluster initialization and limited discrimination in high-dimensional traffic data (Ring & Wunderlich, 2020; Patil & Thorat, 2019). Isolation Forest performed moderately, with 85.2% accuracy, but like k-Means, it suffered from higher false alarms, which reduce trust in operational settings (Latah & Toker, 2021).

Overall, the results confirm that supervised models are superior for detecting known attacks given sufficient labeled data, while unsupervised models offer complementary strength for zero-day detection. However, both paradigms exhibit trade-offs between accuracy, computational cost, and false alarms. These findings highlight the importance of hybrid IDS approaches that combine the predictive accuracy of supervised learning with the anomaly detection capability of unsupervised models (Islam & Abawajy, 2020; Kim et al., 2014).

4.3 Comparative Summary

Model	Accuracy	FPR	AUC
Random Forest	97.3%	2.1%	0.98
SVM	95.5%	3.5%	0.94
Autoencoder	88.4%	6.1%	0.91
k-Means Clustering	83.9%	8.3%	0.89

Isolation Forest	85.2%	7.6%	0.88

5.Discussion

The findings of this research reiterate the general thrust of cybersecurity literature: supervised machine learning models, especially ensembles such as Random Forest, provide better detection accuracy and dependability when trained on properly labeled data. Their effectiveness, however, tapers off when subjected to new attacks or lacking adequate labeled data—a real-world scenario in networked environments.

In contrast, unsupervised models perform better for dynamic, real-time detection applications. Their capacity to detect outliers or anomalies without knowing beforehand makes them well-suited for zero-day attack identification. Even though their precision is relatively lower, their adaptive quality and lesser reliance on labeled datasets provide significant value for next-generation Intrusion Detection Systems (IDS).

One of the more exciting avenues for future work is the creation of a hybrid IDS system that combines both paradigms. This would exploit the predictive capability of supervised models for established attack vectors while also exploiting unsupervised anomaly detection to alert on previously unseen or changing threats. Further, reinforcement learning and semi-supervised methodologies can be used to heighten the system's responsiveness in real-time environments.

Lastly, given the increasing complexity of network infrastructures and the sophistication of threats from malicious individuals, machine learning IDS offerings need to keep updating continuously. Ongoing retraining with fresh data, adversarial testing for robustness, and explainable AI (XAI) incorporation are top areas that need to be researched thoroughly to maintain a long-term and secure cyber environment.

6. Conclusion

This research underscores the potential of machine learning to revolutionize the effectiveness of intrusion detection systems (IDS) in the face of increasingly sophisticated and dynamic cyber threats. Our comparative analysis of supervised and unsupervised ML models, employing benchmark datasets NSL-KDD and CICIDS2017, illustrates that supervised algorithms such as Random Forest and Support Vector Machine attain high accuracy and low rates of false positives in identifying known threats but are highly dependent upon the presence of labeled training data. Conversely, unsupervised models like Autoencoders and

k-Means Clustering are more flexible and are optimized for the identification of never-before-seen or zero-day attacks but with relatively lower accuracy and higher false positives.

Such results highlight the requirement for a hybrid IDS architecture that unites the virtues of both learning paradigms. This type of framework would be able to provide high accuracy in detection against known attacks while still being adaptive enough to find novel patterns of attack without labeling. Merging anomaly-based and signature-based detection into a single, smart system is capable of providing more proactive, adaptive, and context-sensitive cybersecurity defenses.

Possible directions for future work involve analyzing deep learning methods—i.e., Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformers—that provide the potential to represent intricate temporal and spatial dependencies between network traffic data. Further, deploying real-time detection mechanisms in operational environments, model inference speed optimization, and minimizing computational cost without compromising performance continue to be of prime importance for deployment feasibility. Further, explainable AI (XAI) must be implemented to increase interpretability and trust in ML-based IDS solutions, particularly in high-stakes areas such as finance, defense, and healthcare.

Finally, this research offers proof for the integration of machine learning methods into cybersecurity infrastructure, calling for an era where security systems not only react but also predict, dynamically adjust, and learn.

References

Aburomman, A. A., & Reaz, M. B. I. (2016). A survey of machine learning techniques for intrusion detection systems. Journal of Network and Computer Applications, 60, 157–180. https://doi.org/10.1016/j.jnca.2015.11.016

Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

Ali, I., Usman, M., & Shahzad, A. (2020). Detection of cyber attacks using machine learning algorithms. International Journal of Advanced Computer Science and Applications, 11(5), 146–152. https://doi.org/10.14569/IJACSA.2020.0110519

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153–1176. https://doi.org/10.1109/COMST.2015.2494502

Choudhary, A., Kesswani, N., & Sharma, M. (2020). Performance analysis of classification algorithms on NSL-KDD dataset. International Journal of Computer Sciences and Engineering, 8(3), 1–5. https://doi.org/10.26438/ijcse/v8i3.1-5

Dhanabal, L., & Shantharajah, S. P. (2015). A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. International Journal of Advanced Research in Computer and Communication Engineering, 4(6), 446–452. https://doi.org/10.17148/IJARCCE.2015.46205

Gharib, T. F., Alazab, M., Hobbs, M., Abawajy, J., & Lin, X. (2016). Network intrusion detection system using neural networks: A performance evaluation. Procedia Computer Science, 83, 124–130. https://doi.org/10.1016/j.procs.2016.04.105

Islam, R., & Abawajy, J. H. (2020). A hybrid intrusion detection system using supervised and unsupervised machine learning techniques. Journal of Computer and System Sciences, 110, 29–43. https://doi.org/10.1016/j.jcss.2019.10.003

Ring, M., Wunderlich, S., Scheel, C., Landes, D., & Hotho, A. (2019). A survey of network-based intrusion detection data sets. Computers & Security, 86, 147–167. https://doi.org/10.1016/j.cose.2019.06.005

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In Proceedings of the 4th

International Conference on Information Systems Security and Privacy (ICISSP) (pp. 108–116). https://doi.org/10.5220/0006639801080116

Siddiqui, S. T., & Naeem, A. (2021). A deep learning-based framework for anomaly detection in network traffic. Computers, Materials & Continua, 68(1), 587–602. https://doi.org/10.32604/cmc.2021.013240

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In 2010 IEEE Symposium on Security and Privacy (pp. 305–316). IEEE. https://doi.org/10.1109/SP.2010.25

Zhang, J., Zulkernine, M., & Haque, A. (2008). Random-forest-based network intrusion detection systems. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 38(5), 649–659. https://doi.org/10.1109/TSMCC.2008.923876

Shaukat, K., Luo, S., Farooq, M., & Bhatti, F. T. (2020). A survey on machine learning techniques for cyber security in the last decade. Journal of Network and Computer Applications, 101, 183-213.

Ring, M., & Wunderlich, S. (2020). Unsupervised machine learning approaches in network anomaly detection: A survey. International Journal of Network Management, 30(3), e2093.

Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.

Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer Networks, 51(12), 3448-3470.

Li, Y., & Li, M. (2017). Cyber security anomaly detection based on machine learning: A survey. Journal of Security and Communication Networks, 2017, Article ID 2954285.

Buczak, A. L., & Guven, E. (2020). An introduction to data mining and machine learning for cyber security. In Advances in Data Mining for Cybersecurity (pp. 3-20). Springer.

Patil, A., & Thorat, S. S. (2019). Anomaly based intrusion detection system using machine learning. International Journal of Computer Applications, 176(37), 5-8.

Gu, G., Peralta, R. C., & Fu, S. (2016). Malware classification using ontologies and machine learning. International Journal of Information Security Science, 5(4), 121-130.

Mukkamala, S., Janoski, G., & Sung, A. H. (2002). Intrusion detection using neural networks and support vector machines. In Proceedings of the 2002 International Joint Conference on Neural Networks (Vol. 2, pp. 1702-1707). IEEE.

Cannady, J. (1998). Artificial neural networks for misuse detection. Proceedings of the 1998 National Information Systems Security Conference.

Le, N., & Bassett, C. (2018). Machine learning algorithms for cyber security: A systematic review. Journal of Cyber Security and Mobility, 7(1), 61-81.

Latah, M., & Toker, L. (2021). Anomaly detection in cyber security using deep learning: A survey. Journal of Network and Computer Applications, 183, 103070.

Alazab, M., Venkatraman, S., Watters, P., & Layton, R. (2012). Classification of malicious executables using multiple API sequence mining. Journal of Network and Computer Applications, 36(1), 324-335.

Sundaram, S. B., & Ramamurthy, B. (2020). Anomaly detection for intrusion detection using deep learning techniques. Computer Communications, 173, 161-173.

Beghdad, R. (2020). Hybrid genetic algorithm and deep neural network for network intrusion detection system. Neurocomputing, 408, 45-55.

Wang, W., Zhu, M., Zeng, X., Ye, X., & Sheng, Y. (2017). Malware traffic classification using convolutional neural network for representation learning. IEEE Transactions on Network and Service Management, 15(2), 757-770.

Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Applying deep learning approaches for network traffic classification. In IEEE International Conference on Advances in Computing, Communications and Informatics (pp. 1222-1228).

Yang, J., & Liu, Y. (2021). Machine learning for cybersecurity in cloud computing: A comprehensive survey. Security and Communication Networks, 2021, Article ID 9951387.

Zhang, Y., Li, L., & Li, Y. (2020). A survey of machine learning methods applied to intrusion detection. In 2020 12th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC) (Vol. 1, pp. 30-35).

Gaurav, A., & Bhardwaj, A. (2017). A comparative study of classification algorithms for network intrusion detection. International Journal of Advanced Research in Computer Science, 8(7).

Das, S., & Sharma, V. (2018). Intrusion detection systems using machine learning: A comparative review. Procedia Computer Science, 132, 1235-1242.

Conti, M., Dehghantanha, A., Franke, K., & Watson, S. (2018). Internet of Things security and forensics: Challenges and opportunities. Future Generation Computer Systems, 78, 544-546.

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. IEEE Symposium on Security and Privacy, 305-316.

Hodo, E., Bellekens, X., Hamilton, A., Tachtatzis, C., & Atkinson, R. (2017). Threat analysis of IoT networks using artificial neural network intrusion detection system. In ICC 2017 - IEEE International Conference on Communications.

Depren, O., Topallar, M., Anarim, E., & Ciliz, M. (2005). An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks. Expert Systems with Applications, 29(4), 713-722.

Vidyarthi, A., Rathore, S., & Saxena, A. (2019). Machine learning based anomaly detection for network traffic analysis. Journal of Information Security and Applications, 46, 58-67.

Nguyen, T. T., & Armitage, G. (2008). A survey of techniques for internet traffic classification using machine learning. IEEE Communications Surveys & Tutorials, 10(4), 56-76.