

Intelligent Intrusion Detection Systems Using Hybrid Classification and Anomaly Detection Models

Sofia Bellini

Associate Professor, European College of Cybersecurity, Italy

Abstract

Intrusion detection systems analyze network or system activity to identify behavior that may indicate unauthorized or abnormal activity. Classification-based approaches can recognize patterns represented within labeled training data, whereas anomaly-detection methods attempt to identify observations that deviate from established normal behavior. Each approach has methodological limitations when used independently. Supervised classifiers may perform strongly on familiar patterns but depend on representative labels, while anomaly detectors may identify unusual activity without knowing whether every deviation is genuinely security relevant. This hypothetical methodological study examines an intelligent hybrid intrusion detection framework that combines supervised classification with anomaly detection. A fictional dataset containing 5,000 network-session observations is divided chronologically into training, validation, and testing partitions. Three conceptual systems are compared: random-forest classification, one-class support vector anomaly detection, and a hybrid architecture combining classification confidence with anomaly scoring. The study is limited to defensive analytical evaluation and provides no operational offensive procedures.

The illustrative analysis evaluates accuracy, precision, recall, and F1 score using a hypothetical 800-observation test set containing 200 intrusion-labeled and 600 normal observations. Under the assigned values, the hybrid model records the strongest overall classification performance. A second comparison examines detection behavior across familiar intrusion patterns, previously unseen anomalous patterns, low-frequency suspicious behavior, and mixed traffic conditions. Performance varies across these conditions, demonstrating the importance of evaluating both known and unfamiliar events. The study concludes that combining supervised classification and anomaly detection provides a useful methodological framework for intelligent intrusion detection because the two approaches address complementary detection requirements. However, hybridization does not automatically guarantee reliable security performance. Practical effectiveness depends on representative traffic data, appropriate anomaly thresholds, control of false positives, temporally realistic testing, and continuous evaluation as network behavior changes.

Keywords: Intrusion Detection System, Hybrid Classification, Anomaly Detection, Machine Learning, Network Security, Random Forest, One-Class Support Vector Machine, Cybersecurity Analytics, Network Anomaly Detection, Intelligent Security Systems.

1. Introduction

Networked information systems generate large volumes of communication activity that may contain both legitimate and security-relevant behavior. Intrusion detection systems provide an additional defensive layer by monitoring events and identifying patterns that warrant investigation. Detection is challenging because ordinary network activity can vary substantially across users, services, time periods, and infrastructure conditions. An effective IDS must therefore distinguish potentially harmful deviations from legitimate variation without producing an excessive volume of false alarms.

Intrusion detection approaches have traditionally been divided into misuse or signature-based detection and anomaly-based detection. Signature-oriented methods identify activity that corresponds to previously characterized patterns, providing relatively precise recognition when known indicators are available. Anomaly detection instead establishes some representation of expected behavior and identifies significant deviations. Patcha and Park and García-Teodoro et al. emphasized that anomaly-oriented methods can potentially detect unfamiliar behavior, although false-positive control remains a major practical challenge.

Machine learning introduced additional possibilities for intrusion detection by learning decision boundaries directly from network or system data. Lee and Stolfo investigated data-mining approaches for constructing intrusion-detection models, while Mukkamala et al. applied neural networks and support vector machines to intrusion detection. Supervised classification can identify complex feature relationships when sufficiently representative labeled data are available. However, its performance on previously unseen activity may be weaker when novel observations differ substantially from the training distribution.

Unsupervised and one-class approaches address a different detection problem by learning characteristics of normal or unlabeled traffic and identifying atypical observations. Laskov et al. explicitly compared supervised and unsupervised learning for intrusion detection, highlighting differences in their ability to identify known and unknown patterns. Isolation Forest and one-class support vector methods provide general anomaly-detection strategies that can be adapted to security analytics. Nevertheless, abnormality alone does not prove malicious activity, and operational context remains necessary.

The present hypothetical study combines supervised classification and anomaly detection within one conceptual IDS. A random-forest classifier estimates whether network observations resemble labeled intrusion patterns, while a one-class anomaly component evaluates deviations from established normal behavior. A hybrid decision stage integrates both signals. The study compares the three configurations through transparent hypothetical metrics and evaluates the hybrid model across several traffic conditions. No real network environment was attacked, monitored, or tested for this manuscript.

2. Objectives of the Study

The principal objective of this study is to examine the methodological usefulness of combining supervised classification and anomaly detection for intelligent intrusion detection. Particular attention is given to complementary detection behavior, overall classification performance, unfamiliar-pattern recognition, false-positive control, and methodological requirements for reliable cybersecurity evaluation.

2.1 Examination of Hybrid Intrusion Detection Architecture

The first objective is to examine how a supervised classifier and an anomaly detector can be integrated within a common defensive analytical framework. The classifier estimates whether an observation resembles labeled intrusion patterns, whereas the anomaly component estimates deviation from modeled normal behavior. Their outputs are combined through a conceptual decision layer. The hybrid design is intended to retain the precision of classification while providing additional sensitivity to observations not represented clearly in the labeled intrusion classes.

Hybridization is not treated as a simple rule that every anomaly should automatically be classified as an intrusion. Legitimate network changes can also produce unusual patterns. The conceptual framework therefore uses both classifier confidence and anomaly score before generating a higher-priority security alert. This objective emphasizes complementary evidence rather than unconditional combination. A useful hybrid IDS should reduce missed unusual events without creating an unmanageable increase in false-positive alerts.

2.2 Comparison of Detection Performance

The second objective is to compare a random-forest classifier, a one-class support vector anomaly detector, and the integrated hybrid model using identical hypothetical test observations. Accuracy, precision, recall, and F1 score are calculated from consistent confusion matrices. The positive class represents observations assigned an intrusion label within the fictional dataset. This comparison demonstrates how different detection philosophies can be evaluated under the same class prevalence and testing period.

The analysis also examines hybrid-model performance under distinct traffic conditions. Familiar patterns are expected to be handled effectively by the supervised component, whereas unseen or infrequent abnormal patterns may rely more strongly on the anomaly component. Mixed traffic conditions present an additional challenge because normal and suspicious observations can overlap statistically. Evaluating these conditions separately helps determine whether strong aggregate accuracy conceals weakness in the types of events for which anomaly detection was introduced.

2.3 Identification of IDS Deployment Challenges

The third objective is to identify methodological challenges affecting intelligent intrusion detection. Relevant concerns include evolving network behavior, imbalanced attack labels, incomplete ground truth, false-positive alerts, anomaly-threshold selection, dataset aging, and differences between laboratory and operational environments. A model that performs strongly on one benchmark may not maintain equivalent performance when deployed on a network with different services, user behavior, and traffic distributions.

The study further emphasizes the importance of defensive human oversight. An IDS alert indicates that an observation should be investigated; it does not independently establish intent or confirm that a security incident occurred. Analysts require supporting context such as affected systems, timing, confidence, and related events. Intelligent IDS research should therefore evaluate whether automated detection improves investigation quality and prioritization rather than treating algorithmic classification as a substitute for security judgment.

3. Review of Literature

3.1 Machine Learning for Intrusion Classification

Lee and Stolfo developed a data-mining framework for constructing intrusion-detection features and models from audit information. Their work helped establish supervised learning as an important approach to automated security analysis. Mukkamala et al. later evaluated neural networks and support vector machines for intrusion detection, demonstrating how nonlinear classifiers could distinguish normal and suspicious patterns. These contributions supported broader interest in learning-based IDS architectures beyond manually specified security rules.

Peddabachigari et al. investigated hybrid intelligent systems combining decision trees and support vector machines. Their work demonstrated that multiple classifiers could be arranged hierarchically or as ensembles to improve detection performance. Tsai et al. subsequently reviewed machine-learning techniques for intrusion detection and examined single, hybrid, and ensemble approaches across prior studies. The literature indicates that classifier combination can provide useful complementary behavior, although reported performance depends strongly on datasets and evaluation procedures.

3.2 Anomaly Detection for Network Security

Patcha and Park reviewed anomaly-detection methods and emphasized the ability of anomaly-oriented systems to detect behavior not represented explicitly through attack signatures. García-Teodoro et al. similarly reviewed anomaly-based network intrusion detection and identified false positives, adaptability, and assessment methodology as continuing challenges. These studies demonstrate why anomaly detection remains attractive for changing security environments while also explaining why operational adoption requires careful threshold and context management.

Kruegel and Vigna applied anomaly detection to web-based activity, demonstrating how statistical profiles could identify deviations associated with suspicious behavior. Laskov et al. compared supervised and unsupervised intrusion-detection learning directly and examined detection accuracy for known and unknown events. One-class learning and unsupervised anomaly techniques therefore provide a methodological complement to multiclass or binary supervised classification, particularly when complete labeled attack data cannot be assumed.

3.3 Hybrid IDS and Dataset Evaluation

Aydin et al. developed a hybrid intrusion detection system combining misuse-oriented and anomaly-oriented detection mechanisms. Their approach illustrates the central logic of hybrid IDS design: known patterns can be identified through one component while anomaly methods provide additional coverage for behavior that does not match established signatures. Hybrid systems therefore attempt to reduce the individual limitations of purely misuse-based or purely anomaly-based detection.

Evaluation datasets remain equally important. Tavallaei et al. identified limitations within KDD CUP 99 and proposed NSL-KDD to reduce redundancy-related evaluation problems. Moustafa and Slay later introduced UNSW-NB15 to represent a more contemporary mixture of normal and synthetic malicious traffic. These contributions demonstrate that IDS performance is inseparable from dataset composition. High benchmark accuracy does not establish operational reliability when traffic characteristics differ substantially from real deployment conditions.

4. Research Methodology

4.1 Research Design and Hybrid Detection Framework

The present study adopts a hypothetical comparative methodological design involving three defensive detection configurations. The first is a random-forest supervised classifier trained on labeled normal and intrusion observations. The second is a one-class support vector machine trained primarily on normal traffic to identify statistical deviations. The third integrates classifier probability and standardized anomaly score through a conceptual hybrid decision function. No model was actually trained or deployed for this manuscript.

The hybrid model first evaluates whether an observation corresponds strongly to a labeled intrusion pattern. Observations with lower classifier certainty are additionally examined according to anomaly score. High-confidence normal observations remain lower priority, whereas observations receiving either strong intrusion-class evidence or combined moderate classification and anomaly evidence are elevated for review. The framework is intentionally described at an analytical level and does not include procedures for exploiting, bypassing, or attacking real systems.

4.2 Hypothetical Network Dataset and Features

The fictional dataset contains 5,000 network-session observations represented through twenty-four defensive traffic features. Three thousand five hundred observations are allocated chronologically to training, seven hundred to validation, and eight hundred to independent testing. The test partition contains 200 intrusion-labeled and 600 normal observations. All traffic records, addresses, labels, and network events are fictional and do not correspond to any operational organization or public benchmark dataset.

Features conceptually include connection duration, packet-count summaries, directional byte totals, protocol category, connection-state characteristics, timing variation, connection-frequency measures, and normalized service-level behavior. Direct identifying addresses are excluded. Feature scaling parameters and anomaly thresholds are conceptually determined using training and validation information only. Future test observations are not used during model preparation, thereby reducing the risk of temporal leakage and overly optimistic evaluation.

4.3 Evaluation Measures and Analytical Procedure

Detection performance is evaluated through accuracy, precision, recall, and F1 score. True positives represent intrusion-labeled observations correctly detected, while false negatives represent missed intrusion observations. False positives represent normal observations incorrectly flagged, and true negatives represent correctly classified normal traffic. Precision is especially relevant to analyst workload, while recall indicates the ability to identify the intrusion observations contained in the hypothetical test set.

The second analysis evaluates the hybrid model under four conceptual traffic conditions: familiar intrusion patterns, previously unseen anomalous patterns, low-frequency suspicious patterns, and mixed high-variability traffic. Detection rate, false-positive rate, and F1 score are reported for illustration. These categories do not describe operational attack procedures. A real evaluation would additionally require repeated temporal tests, calibration analysis, alert-volume assessment, and external validation using independent network environments.

5. Analysis

5.1 Comparative Detection Performance

The first hypothetical analysis compares the random-forest classifier, one-class anomaly detector, and hybrid model using the same 800-observation testing period. The supervised classifier correctly identifies 168 of 200 intrusion observations, while the anomaly model identifies 158. The hybrid configuration detects 184 intrusion-labeled observations and generates thirty-four false positives among 600 normal observations. The assigned results illustrate how combining complementary signals could reduce missed events without allowing anomaly alerts to dominate the decision process.

The hybrid system records hypothetical accuracy of 93.75%, precision of 84.40%, recall of 92.00%, and F1 score of 88.04%. The random forest records an F1 score of 81.36%, while the one-class SVM records 73.83%. These results are deliberately constructed and do not establish that hybrid IDS models universally outperform individual techniques. Empirical conclusions would require real training, repeated experiments, optimized baselines, and confidence intervals.

Table 1: Hypothetical Comparative Performance of Intrusion Detection Models

Detection Model	TP	FP	FN	TN	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Random Forest Classification	168	45	32	555	90.38	78.87	84.00	81.36
One-Class SVM Anomaly Detection	158	70	42	530	86.00	69.30	79.00	73.83
Hybrid Classification + Anomaly Model	184	34	16	566	93.75	84.40	92.00	88.04

Note: All observations and performance values are fictional and were not produced by an operational intrusion detection system.

5.2 Hybrid Detection Across Traffic Conditions

The second hypothetical analysis examines the hybrid IDS under four high-level testing conditions. Familiar intrusion patterns produce the strongest assigned F1 score because supervised classification can use patterns represented clearly during training. Previously unseen anomalous patterns show somewhat lower precision but maintain comparatively strong detection because the anomaly component contributes additional evidence. Low-frequency suspicious observations are more difficult because limited examples provide weaker statistical separation from unusual but legitimate traffic.

Mixed high-variability traffic produces the largest illustrative false-positive rate because changing legitimate behavior increases the number of normal observations that differ from the training distribution. This pattern demonstrates an important operational concern: anomaly detection becomes more difficult when the definition of normal traffic is itself unstable. The assigned values are methodological examples only and should not be interpreted as measured performance against any real intrusion category.

Table 2: Hypothetical Hybrid IDS Performance Across Traffic Conditions

Traffic Condition	Detection Rate (%)	False-Positive Rate (%)	F1 Score (%)
Familiar Intrusion Patterns	95.0	3.8	91.8
Previously Unseen Anomalous Patterns	89.0	5.6	85.2
Low-Frequency Suspicious Patterns	84.0	6.5	80.7
Mixed High-Variability Traffic	87.0	8.3	81.9

Note: Traffic-condition labels and numerical values are fictional and do not describe attacks against a real network.

6. Discussion

6.1 Complementarity of Classification and Anomaly Detection

The hypothetical findings illustrate the principal rationale for hybrid intrusion detection. Supervised classifiers can learn detailed decision boundaries from labeled examples and therefore provide strong recognition of familiar intrusion patterns. Anomaly detection addresses a complementary problem by identifying behavior that differs from learned normality even when no explicit intrusion label exists. Laskov et al. demonstrated the methodological distinction between supervised and unsupervised intrusion detection, while Aydin et al. showed the potential value of combining detection paradigms.

Hybrid performance nevertheless depends on how component outputs are integrated. An anomaly detector with an excessively sensitive threshold can overwhelm a precise supervised classifier with false alerts, whereas an overly restrictive threshold may provide little additional coverage. The decision stage therefore requires validation against realistic alert costs and class prevalence. Hybridization should be justified by measurable improvement in security-relevant outcomes rather than by architectural complexity alone.

6.2 False Positives and Operational Reliability

False positives remain one of the most important practical limitations of anomaly-based intrusion detection. Axelsson demonstrated that low event prevalence can create substantial difficulties even when detection sensitivity appears strong. A relatively small false-positive rate can still generate many alerts when normal traffic vastly outnumbers genuine security incidents. Precision and alert volume are therefore essential evaluation measures alongside detection rate.

The hypothetical hybrid model reduces false positives relative to the standalone anomaly detector because anomaly evidence is interpreted together with classification confidence. This illustrates one possible benefit of combining methods. However, actual performance would depend on traffic diversity and threshold calibration. Analysts should evaluate how many alerts are generated per operational period, whether alerts contain useful context, and whether the detection system supports efficient investigation rather than simply maximizing benchmark classification metrics.

6.3 Dataset Quality and Generalization

Intrusion-detection models can appear highly effective when training and testing datasets contain redundant or closely related observations. Tavallaei et al. demonstrated how characteristics of KDD CUP 99 could distort evaluation and motivated the development of NSL-KDD. Moustafa and Slay introduced UNSW-NB15 partly in response to the need for network-security datasets reflecting more contemporary traffic characteristics. These developments demonstrate that benchmark design directly influences conclusions about IDS performance.

Sommer and Paxson further emphasized the difficulty of transferring machine-learning results from closed benchmark settings to operational network environments. Real traffic distributions evolve, labels are incomplete, and the cost of errors can differ from laboratory assumptions. A robust intelligent IDS should therefore be evaluated through temporally separated data, independent environments, and changing normal behavior. Generalization is a more meaningful objective than maximizing performance on one fixed dataset.

7. Challenges

7.1 Changing Network Behavior and Concept Drift

Normal network behavior changes as organizations introduce new applications, modify infrastructure, alter user workflows, or experience changes in traffic volume. An anomaly model trained on an earlier period may therefore begin flagging legitimate new behavior as suspicious. This phenomenon can increase false-positive rates even when the model itself has not changed. Periodic reassessment is necessary to determine whether the learned normal profile remains representative.

Updating the model introduces a related risk because new training data may contain undetected suspicious observations or temporary anomalies. Automatic retraining should therefore not occur without appropriate validation. Model versions, training periods, thresholds, and data-quality checks should be documented. Adaptation is valuable only when it improves representation of legitimate network evolution without weakening the system's ability to identify genuinely unusual activity.

7.2 Class Imbalance and Incomplete Security Labels

Intrusion datasets commonly contain substantially more normal observations than verified malicious events. This imbalance can make overall accuracy misleading because a classifier may perform strongly simply by favoring the majority class. The present analysis therefore reports precision, recall, and F1 score in addition to accuracy. Real evaluations may also require precision-recall curves, class-sensitive thresholds, or cost-weighted measures depending on operational priorities.

Ground-truth labels present another difficulty. Security events may be missed, incorrectly categorized, or discovered only after investigation. Supervised models inherit weaknesses in these labels. Anomaly detection can reduce dependence on complete attack labeling but cannot determine automatically whether every unusual event is harmful. Intelligent intrusion detection therefore requires an interaction between statistical detection and analyst-confirmed security context.

7.3 Hybrid Complexity and Explainability

Hybrid systems contain multiple models, thresholds, and decision rules, increasing the complexity of determining why a particular alert was produced. Security analysts may need to know whether an observation was flagged because it resembled a known intrusion pattern, because its behavior was

statistically unusual, or because both conditions occurred. This distinction can affect investigation priority and confidence.

Explainability does not require disclosure of every mathematical detail to an operator, but the system should provide understandable alert context. Relevant features, anomaly magnitude, classifier confidence, and historical comparison can support interpretation. Model complexity should be increased only when it provides measurable detection benefit. A simpler system with slightly lower benchmark accuracy may be preferable if it is substantially easier to validate, maintain, and interpret in an operational environment.

8. Conclusion

Intelligent intrusion detection can benefit from combining supervised classification and anomaly detection because the two methods address different aspects of network-security monitoring. The present hypothetical study compared random-forest classification, one-class anomaly detection, and an integrated hybrid model within a common evaluation framework. The methodology emphasized both detection of familiar labeled patterns and identification of unusual observations that may not correspond directly to known classes.

The illustrative analysis assigned the highest accuracy, recall, and F1 score to the hybrid configuration and showed that detection behavior varied across familiar, previously unseen, low-frequency, and highly variable traffic conditions. Every numerical result is fictional and does not constitute evidence from a deployed cybersecurity system. Real claims would require validated network data, repeatable experiments, temporal testing, threshold calibration, and independent operational evaluation.

The study concludes that hybrid IDS architectures are methodologically promising but should be evaluated primarily through operationally meaningful criteria, including false-positive burden, unknown-pattern sensitivity, generalization, and analyst usability. Future empirical work should examine adaptive anomaly thresholds, concept drift, imbalanced labels, explainable alerts, and cross-network validation. Intelligent detection should support defensive investigation rather than replace security analysis or incident verification.

References

1. Ali, A. H., Charfeddine, M., Ammar, B., Ben Hamed, B., Albalwy, F., Alqarafi, A., & Hussain, A. (2024). Unveiling machine learning strategies and considerations in intrusion detection systems: A comprehensive survey. *Frontiers in Computer Science*, 6, Article 1387354. <https://doi.org/10.3389/fcomp.2024.1387354>
2. Al-Bander, B., Yusof, R., Saif, A., et al. (2024). Meta-analysis and systematic review for anomaly network intrusion detection systems: Detection methods, dataset, validation methodology, and challenges. *IET Networks*, 13(5–6), 339–376. <https://doi.org/10.1049/ntw2.12128>
3. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
4. Denning, D. E. (1987). An intrusion-detection model. *IEEE Transactions on Software Engineering*, SE-13(2), 222–232. <https://doi.org/10.1109/TSE.1987.232894>
5. Dua, D., & Graff, C. (2017). UCI machine learning repository. University of California, Irvine, School of Information and Computer Sciences. <https://archive.ics.uci.edu/>

6. Eskin, E., Arnold, A., Prerau, M., Portnoy, L., & Stolfo, S. (2002). A geometric framework for unsupervised anomaly detection: Detecting intrusions in unlabeled data. In *Applications of data mining in computer security* (pp. 77–101). Springer. https://doi.org/10.1007/3-540-34471-0_4
7. Garcia-Teodoro, P., Díaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security*, 28(1–2), 18–28. <https://doi.org/10.1016/j.cose.2008.08.003>
8. Hodge, V. J., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22, 85–126. <https://doi.org/10.1023/B:AIRE.0000045502.10941.a9>
9. Kimanzi, R., Kimanga, P., Cherori, D., & Gikunda, P. K. (2024). Deep learning algorithms used in intrusion detection systems—A review. *arXiv*. <https://arxiv.org/abs/2402.17020>
10. Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. *Cybersecurity*, 2, Article 20. <https://doi.org/10.1186/s42400-019-0038-7>
11. Laskov, P., Düssel, P., Schäfer, C., & Rieck, K. (2005). Learning intrusion detection: Supervised or unsupervised? In *Proceedings of the 13th International Conference on Image Analysis and Processing* (pp. 50–57). IEEE. <https://doi.org/10.1109/ICIAP.2005.103>
12. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In *2015 Military Communications and Information Systems Conference (MilCIS)* (pp. 1–6). IEEE. <https://doi.org/10.1109/MilCIS.2015.7348942>
13. Ruffo, V. G. da S., Lent, D. M. B., Komarchesqui, M., Schiavon, V. F., de Assis, M. V. O., Carvalho, L. F., & Proença, M. L., Jr. (2024). Anomaly and intrusion detection using deep learning for software-defined networks: A survey. *Expert Systems with Applications*, 256, Article 124982. <https://doi.org/10.1016/j.eswa.2024.124982>
14. Sajid, M., Malik, K. R., Almogren, A., Malik, T. S., Khan, A. H., Tanveer, J., & Ur Rehman, A. (2024). Enhancing intrusion detection: A hybrid machine and deep learning approach. *Journal of Cloud Computing*, 13, Article 123. <https://doi.org/10.1186/s13677-024-00685-x>
15. Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7), 1443–1471. <https://doi.org/10.1162/089976601750264965>
16. 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* (pp. 108–116). SciTePress. <https://doi.org/10.5220/0006639801080116>
17. 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>
18. Tavallaei, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. In *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications* (pp. 1–6). IEEE. <https://doi.org/10.1109/CISDA.2009.5356528>
19. Thakkar, A., & Lohiya, R. (2021). A review on machine learning and deep learning perspectives of IDS for IoT: Recent updates, security issues, and challenges. *Archives of Computational Methods in Engineering*, 28, 3211–3243. <https://doi.org/10.1007/s11831-020-09496-0>
20. Landscape of learning techniques for intrusion detection system in IoT: A systematic literature review. (2024). *Computers & Electrical Engineering*, 120(Part B), Article 109725. <https://doi.org/10.1016/j.compeleceng.2024.109725>