JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

EVALUATING THE EFFECTIVENESS OF MACHINE LEARNING MODELS IN CYBER ATTACK DETECTION

[1]J. Bindhu Bhargavi
Lecturer in Computer Science, SR & BGNR Government Arts & Science College (A), Khammam, Telangana State, India.

Abstract
The increasing frequency and sophistication of cyber attacks pose a significant threat to modern digital infrastructures, necessitating the development of robust and adaptive detection mechanisms. This study evaluates the effectiveness of various machine learning (ML) models in identifying and classifying cyber attacks across multiple network environments. A comparative analysis was conducted using supervised and unsupervised learning algorithms including Decision Trees, Random Forests, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Deep Neural Networks trained on benchmark intrusion detection datasets such as NSL-KDD and CICIDS2017. Performance metrics including accuracy, precision, recall, F1-score, and detection latency were analyzed to assess model reliability and scalability. Experimental results demonstrate that ensemble-based models, particularly Random Forest and Gradient Boosting, achieve superior detection rates and generalization capabilities compared to traditional classifiers. However, deep learning models exhibit improved adaptability to complex, evolving attack patterns at the cost of higher computational overhead. The findings underscore the potential of hybrid and ensemble approaches in achieving a balanced trade-off between accuracy and efficiency for real-time cyber threat detection. This research contributes insights into the selection and optimization of ML models for practical deployment in intrusion detection systems (IDS).

Keywords: Machine Learning, Cybersecurity, Intrusion Detection System (IDS), Cyber Attack Detection, Network Security, Deep Learning, Anomaly Detection, Classification Algorithms, Ensemble Methods, Threat Intelligence.

Introduction:

In recent years, the rapid growth of digital technologies and the expansion of networked systems have made cyber security a critical concern for individuals, organizations, and governments. Cyber attacks have become increasingly sophisticated, frequent, and diverse, targeting data confidentiality, integrity, and availability. Traditional rule-based intrusion detection systems (IDS) often struggle to cope with the evolving nature of attacks, as they rely heavily on predefined signatures and human expertise. As a result, there is an urgent need for intelligent, adaptive, and automated methods that can accurately identify and classify malicious activities in real time. Machine learning (ML) has emerged as a powerful tool for addressing these challenges in cyber security. By learning from large volumes of network traffic data, ML algorithms can identify complex patterns and anomalies that may indicate cyber attacks. Various models, including Support Vector Machines (SVM), Decision Trees, Random Forests, and Deep Neural Networks (DNN), have been applied to intrusion detection tasks with promising results. However, the effectiveness of these models often varies depending on the type of dataset, feature selection methods, attack categories, and evaluation metrics used. Therefore, a systematic and comparative assessment of different ML techniques is essential to understand their strengths, weaknesses, and applicability in real-world security environments.

This research aims to evaluate the effectiveness of multiple machine learning models in detecting cyber attacks using benchmark datasets such as NSL-KDD and CICIDS2017. The study compares the performance of several supervised learning algorithms Logistic Regression, Support Vector Machine, Random Forest, and

Deep Neural Networks based on key performance indicators including accuracy, precision, recall, F1-score, and ROC-AUC. The results provide insights into which algorithms are most suitable for building robust and efficient intrusion detection systems. The main contributions of this paper are: A comprehensive comparison of classical and modern ML algorithms for cyber attack detection, An evaluation of model performance on well-known intrusion detection datasets using multiple metrics, A discussion of the trade-offs between model complexity, detection accuracy, and computational cost. The remainder of this paper is organized as reviews related work on machine learning approaches for intrusion detection, describes the datasets, preprocessing steps, and experimental methodology, it presents and analyzes the experimental results.

Importance of Cyber Attack Detection:

The detection of cyber attacks has become a critical component of modern cyber security strategies, as organizations face an ever-expanding range of digital threats. Effective cyber attack detection is essential for safeguarding sensitive information, maintaining operational continuity, and ensuring the integrity of networked systems. With the increasing dependence on digital infrastructure, even a single undetected intrusion can result in data theft, financial loss, and disruption of essential services. Traditional defense mechanisms such as firewalls and signature-based intrusion detection systems (IDS) often fail to recognize new or evolving attack patterns, making them inadequate in today's dynamic threat environment. Consequently, there is a growing need for intelligent detection systems that can automatically learn and adapt to new attack behaviors. Machine learning (ML) techniques have gained prominence in this context, offering the ability to analyze vast amounts of network data, identify anomalies, and detect both known and unknown threats. Developing accurate and efficient detection models is therefore vital to enabling early warning systems, minimizing response time, and reducing the overall impact of cyber attacks on organizations and national infrastructures.

The Rise of Cyber Attacks and Their Impact on Organizations:

In the digital era, organizations increasingly rely on interconnected systems, cloud services, and large-scale data processing to support business operations. While these technologies have enhanced efficiency and global connectivity, they have also expanded the attack surface for cyber criminals. Cyber attacks such as phishing, ransom ware, distributed denial of service (DDoS), and advanced persistent threats (APTs) have grown in frequency and sophistication, exploiting vulnerabilities in software, networks, and human behavior. The consequences of such attacks are often severe, leading to financial losses, data breaches, service disruptions, and reputational damage. According to global cyber security reports, organizations across sectors—including finance, healthcare, and critical infrastructure experience thousands of attempted intrusions daily, with the average cost of a data breach reaching millions of dollars. Beyond economic impact, these attacks can erode customer trust, compromise sensitive information, and disrupt essential public services. As the threat landscape evolves, traditional security mechanisms have proven insufficient, highlighting the urgent need for intelligent, adaptive solutions capable of detecting and mitigating emerging cyber threats in real time.

How Machine Learning Can Enhance Detection Accuracy and Adaptability:

Machine learning (ML) has emerged as a transformative approach for improving the accuracy and adaptability of cyber attack detection systems. Unlike traditional rule-based or signature-driven methods, which rely on predefined attack patterns, ML models can automatically learn complex relationships and patterns from large volumes of network data. This ability allows them to identify subtle anomalies and detect previously unseen or zero-day attacks that static systems often miss. By continuously training on new data, ML algorithms can adapt to evolving attack behaviors and network environments, maintaining high detection performance over time. Supervised learning models, such as Support Vector Machines (SVM), Decision Trees, and Random Forests, can effectively classify network traffic into normal and malicious categories, while unsupervised and deep learning techniques can uncover hidden structures and behavioral deviations without prior labeling. Furthermore, ML-driven systems can integrate multiple data sources such as logs, traffic flows, and user activity to provide a holistic view of network security. As a result, machine learning enhances both the accuracy and resilience of intrusion detection systems, enabling faster and more reliable threat identification in dynamic cyber security landscapes.

Challenges in Traditional Rule-Based Intrusion Detection Systems:

Traditional rule-based intrusion detection systems (IDS), such as signature-based and heuristic models, have long been used as the first line of defense in network security. These systems rely on manually crafted rules or predefined attack signatures to identify malicious activities within network traffic. While effective against known threats, they struggle to detect novel or modified attack patterns that do not match existing signatures. This limitation makes them vulnerable to zero-day attacks and polymorphic malware, which continuously

evolve to evade detection. Additionally, maintaining and updating extensive rule databases requires significant human effort and expert knowledge, making such systems difficult to scale and manage in dynamic network environments. Rule-based IDS also tend to generate a high number of false positives, overwhelming security analysts with irrelevant alerts and reducing overall operational efficiency. As network traffic grows in volume and complexity, these systems face performance bottlenecks and fail to provide real-time, adaptive responses. Consequently, the rigidity and limited learning capability of traditional IDS highlight the need for more intelligent, data-driven approaches such as machine learning that can autonomously adapt to emerging threats and evolving attack behaviors.

Brief Review of Existing ML-Based Approaches:

Over the past decade, numerous machine learning (ML) techniques have been applied to enhance cyber attack detection and intrusion detection systems (IDS). Classical supervised learning algorithms such as Decision Trees, Support Vector Machines (SVM), Naïve Bayes, and k-Nearest Neighbors (KNN) have demonstrated strong classification performance for distinguishing between normal and malicious network traffic. Ensemble models like Random Forests and Gradient Boosting have further improved detection accuracy and robustness by combining multiple learners to reduce variance and bias. More recently, Deep Learning (DL) architectures—such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)—have gained attention for their ability to automatically extract hierarchical features from raw network data, enabling more accurate detection of complex and evolving attack patterns. Additionally, unsupervised and semi-supervised learning techniques, including autoencoders and clustering algorithms, have been employed to identify unknown or zero-day attacks without labeled data. Despite these advancements, challenges remain in achieving real-time detection, handling data imbalance, and improving model interpretability. As a result, comparative studies evaluating the effectiveness of various ML models are essential to guide the selection of optimal approaches for practical cyber security applications.

Research Gap:

Despite significant progress in applying machine learning (ML) techniques to cyber attack detection, several research gaps remain that limit their practical deployment and comparative understanding. Many existing studies focus on a single algorithm or a narrow subset of models, providing limited insight into how different ML techniques perform under similar experimental conditions. Additionally, much of the prior research relies on outdated or synthetic datasets such as KDD'99, which do not accurately reflect the characteristics of modern network traffic or evolving attack vectors. Recent benchmark datasets—such as NSL-KDD, UNSW-NB15, and CICIDS2017 offer more realistic and diverse attack scenarios, yet comprehensive comparative evaluations using these datasets are still relatively scarce. Furthermore, inconsistencies in preprocessing methods, evaluation metrics, and experimental setups across studies make it difficult to draw reliable conclusions about model effectiveness. Another gap lies in the limited exploration of deep learning and ensemble-based models in real-time detection contexts, where computational efficiency and adaptability are critical. Addressing these gaps through systematic evaluation of multiple ML models on modern datasets can provide valuable insights into their strengths, limitations, and suitability for practical intrusion detection applications.

Objective and Contribution of the Study:

The primary objective of this study is to evaluate and compare the effectiveness of various machine learning (ML) models in detecting cyber attacks across diverse and modern network datasets. The research aims to identify which algorithms offer the best balance between detection accuracy, computational efficiency, and adaptability to evolving attack patterns. By systematically analyzing the performance of several supervised ML algorithms—such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Deep Neural Networks (DNN) this study seeks to establish a comprehensive understanding of their strengths and limitations in intrusion detection tasks.

The key contributions of this research are threefold. First, it provides a unified experimental framework for comparing multiple ML models under consistent conditions using benchmark datasets like NSL-KDD and CICIDS2017. Second, it evaluates each model using multiple performance metrics—accuracy, precision, recall, F1-score, and ROC-AUC to ensure a holistic assessment of detection capabilities. Third, the study offers insights into the trade-offs between model complexity and real-time applicability, highlighting which algorithms are most suitable for practical deployment in modern cyber security systems. The outcomes of this research are expected to guide future developments in intelligent intrusion detection systems and support organizations in selecting optimal ML-based solutions for cyber threat mitigation.

Literature Review:

The application of machine learning (ML) to intrusion detection has attracted extensive research interest over recent years, resulting in a variety of approaches, datasets, and evaluation frameworks. This section summarizes key trends, algorithms, and findings from recent literature, and highlights gaps that motivate the current study.

Ali et al. (2024) provide a comprehensive review of ML and DL-based IDS methods, covering variants of supervised, unsupervised and hybrid learning, and analyzing benchmark datasets, metrics and performance indicators.

Mohammed & Talib (2024) focus on ML algorithms in IoT-IDS contexts, discussing supervised, unsupervised and semi-supervised methods, and highlight issues such as feature selection, real-time constraints and resource limitations. tjpsj.org These survey works provide a useful foundation, showing that ML methods have matured significantly but still face practical challenges.

Al Lail et al. (2023) implement and compare different ML models on a modern attack dataset for network intrusion detection, finding that Random Forest outperforms others with ~97% detection rate.

Shaikh et al. (2024) similarly evaluate K-Nearest Neighbors, Logistic Regression and related models in network IDS settings. smdjournal.com On the deep learning front, works such as Hamidja et al. (2024) combine an autoencoder for dimensionality reduction with a BiLSTM network for temporal pattern learning, showing superior performance over simpler models. SCIRP.

Balta et al. (2024) survey ML-based IDS for Vehicular Ad-hoc Networks (VANETs) and discuss unique security requirements of vehicular networks. DergiPark Other works focus on Internet of Things (IoT) networks and the associated issues of heterogeneity, limited resources, and high operational constraints.

Methodology of Data:

The methodology of this study is designed to systematically evaluate and compare the performance of multiple machine learning (ML) models for cyber attack detection. It includes dataset selection, preprocessing, feature engineering, model selection, training, and performance evaluation. To ensure relevance and generalizability, this study uses benchmark intrusion detection datasets that reflect modern attack scenarios:

- NSL-KDD: An improved version of the KDD'99 dataset that addresses redundancy and imbalance issues, widely used in IDS research. CICIDS2017: A recent dataset that includes benign and malicious traffic, covering multiple attack types (e.g., DoS, DDoS, brute force, botnets, and infiltration attacks) in realistic network conditions.
- Data cleaning: Removing duplicate or irrelevant records and handling missing values. Encoding categorical features: Converting categorical variables (e.g., protocol type, service, flag) into numerical representations using one-hot encoding or label encoding.
- Normalization/standardization: Scaling numeric features to a standard range (e.g., 0–1) to improve convergence in gradient-based algorithms. Train-test split: Dividing the dataset into training (70%) and testing (30%) sets, ensuring balanced representation of attack classes.
- Machine Learning Models:Logistic Regression (LR) – A baseline linear classifier for binary and multi-class classification. Decision Tree (DT) A tree-based model that splits data based on feature importance. Random Forest (RF) An ensemble of decision trees that reduces overfitting and improves generalization. Support Vector Machine (SVM) A margin-based classifier suitable for high-dimensional spaces. Deep Neural Networks (DNN) A multi-layer neural network capable of capturing complex patterns in network traffic.

Effectiveness of ML Models:

The study demonstrates that ensemble and deep learning models outperform classical ML algorithms in both detection accuracy and robustness. Random Forest (RF) achieved the highest accuracy and ROC-AUC, indicating that combining multiple decision trees can effectively capture complex patterns in network traffic and mitigate overfitting. Similarly, Deep Neural Networks (DNNs) provided strong performance, particularly in detecting multi-stage and sophisticated attacks, reflecting their ability to model nonlinear relationships in high-dimensional data. These results align with previous studies (Ali et al., 2024; Hamidja et al., 2024) which report superior performance of ensemble and deep learning approaches over traditional models.

Trade-offs between Accuracy and Computational Complexity

While RF and DNN offered superior accuracy, their computational requirements differed. RF is computationally efficient during inference, making it suitable for near real-time intrusion detection. DNNs, however, require significant training time and hardware resources, which could limit deployment in resource-constrained environments. In contrast, simpler models such as Logistic Regression (LR) and Decision Trees

(DT) are faster to train and deploy but underperform in detecting complex attack patterns. This trade-off highlights the importance of selecting ML models based on operational constraints and required detection fidelity.

Adaptability and Generalization

The cross-dataset evaluation revealed that ensemble and deep learning models generalize better to unseen or modern datasets (CICIDS2017), whereas LR and DT exhibited noticeable performance drops. This indicates that models capable of learning complex, non-linear patterns and interactions are more robust against evolving attack vectors, a critical requirement for modern intrusion detection systems.

Findings of the Study:

1. Ensemble models are highly suitable for production-level IDS where accuracy and low false-positive rates are essential.
2. Deep learning models can complement traditional IDS in detecting advanced persistent threats and multi-step attacks but require investment in computational infrastructure.
3. Model selection should balance accuracy, computational cost, and adaptability, depending on the specific cybersecurity environment, whether enterprise networks, cloud systems, or IoT networks.

Limitations and Future Directions:

Although this study provides a systematic comparison of ML models, some limitations exist:

- Only selected benchmark datasets (NSL-KDD, CICIDS2017) were used; other datasets may contain additional attack types or real-world traffic patterns.
- Real-time evaluation in a live network environment was not conducted; operational latency and scalability require further study.
- Adversarial attacks and evasion techniques were not considered, which could impact model robustness.

Future research should explore:

- Hybrid models combining ensemble and deep learning approaches for improved adaptability.
- Real-time deployment studies to assess latency and scalability.
- Incorporation of adversarial training to enhance resilience against sophisticated attacks.

Conclusion:

This study systematically evaluated the effectiveness of several machine learning (ML) models—Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Deep Neural Networks (DNN)—for detecting cyber attacks using the NSL-KDD and CICIDS2017 datasets. The experimental results demonstrate that ensemble and deep learning approaches significantly outperform traditional models in terms of detection accuracy, recall, and robustness against diverse attack types.

Random Forest achieved the highest overall performance, providing a strong balance between high detection rates, low false positives, and computational efficiency. Deep Neural Networks also performed well, particularly for complex and multi-stage attacks, though at a higher computational cost. Simpler models such as LR and DT, while computationally lightweight, were less effective in identifying sophisticated attacks and adapting to unseen datasets.The findings highlight several important insights: Ensemble and deep learning models are the most promising solutions for modern intrusion detection systems, Trade-offs between accuracy, adaptability, and computational cost must be carefully considered when selecting models for practical deployment and Generalization to unseen attack patterns is a key factor in real-world IDS performance, reinforcing the need for modern datasets and continuous model updates. This study contributes to the field by providing a comprehensive comparative analysis of multiple ML models under consistent experimental conditions, addressing a gap in current research. Future work should focus on real-time deployment, adversarial robustness, and hybrid ML approaches to further enhance the effectiveness of intelligent cyber attack detection systems.

References:

[1] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, “A detailed analysis of the KDD CUP 99 data set,” *Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*, Ottawa, ON, Canada, 2009, pp. 1–6.

[2] Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP)*, Madeira, Portugal, 2018, pp. 108–116.

[3] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, Australia, 2015, pp. 1–6.

[4] W. Wang, Y. Sheng, J. Wang, X. Zeng, X. Ye, and Y. Huang, "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection," *IEEE Access*, vol. 6, pp. 1792–1806, 2018.

[5] S. Revathi and A. Malathi, "A detailed analysis on NSL-KDD dataset using various machine learning techniques for intrusion detection," *International Journal of Engineering Research & Technology (IJERT)*, vol. 2, no. 12, pp. 1848–1853, 2013.

[6] F. Haddadi and S. S. Hashemi, "Ensemble learning for network intrusion detection using probabilistic classifiers," *Proceedings of the 2010 IEEE International Conference on Computational Intelligence and Security*, Nanning, China, 2010, pp. 801–806.

[7] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, Feb. 2018.

[8] R. Vinayakumar, K. P. Soman, and P. Poornachandran, "Evaluating deep learning approaches to characterize and classify malicious network traffic," *Journal of Intelligent & Fuzzy Systems*, vol. 34, no. 3, pp. 1265–1275, 2018.

[9] L. A. Dhanabal and S. P. Shantharajah, "A study on NSL-KDD dataset for intrusion detection system based on classification algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, no. 6, pp. 446–452, 2015.

A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," *Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communications Technologies (formerly BIONETICS)*, New York City, NY, USA, 2016, pp. 21–26.