Avalokiteshvara Journal of Artificial Intelligence

http://hcapit.org/ajai.html

ISSN: XXXX-XXXX (Pending)

Review Article

Enhancing Network Security: A Comprehensive Review of Deep Learning Models and Datasets for IDS

Padmapani P. Tribhuvan[1], Amrapali P. Tribhuvan[2]

[1]hCAP Institute of Technology, Chhatrapati Sambhajinagar, MS, India.

[2]Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar, MS, India.

padmapani.prakash@gmail.com

amrapaliprakash512@gmail.com

Corresponding Author Email: padmapani.praksh@gmail.com

Pre-publication Copy - Not for Distribution

Abstract

Intrusion Detection Systems (IDS) are indispensable in safeguarding computer networks from increasingly diverse cyber threats. Traditional methods, while effective for known attacks, struggle with the detection of novel and sophisticated threats. Deep Learning (DL) models have emerged as promising to enhance IDS capabilities by automatically learning and extracting complex patterns from network data. This paper comprehensively reviews various DL models applied in IDS, examining their applications, datasets, strengths, challenges, and future research directions.

Keywords: Intrusion Detection System, Deep Learning, Convolutional Neural Network, Re-current Neural Network, Generative Adversarial Network, Long Short-Term Memory

1. Introduction

An Intrusion Detection System (IDS) is a technology designed to identify and react to unauthorized access or malicious activities on computer systems and networks. The main purpose of an IDS is to monitor and analyse network and system activities, detect potential security incidents, and deliver timely alerts or take automated actions to mitigate threats.

IDSs are vital for boosting the overall security of computer systems and networks. They offer a proactive approach to identifying and responding to potential security threats. They are often used together with other security measures, such as firewalls, antivirus software, and security policies, to create secure computer systems and networks.

Traditional Intrusion Detection Systems typically rely on rule-based or signature-based detection methods. They detect known attack patterns by comparing network traffic or system events against a predefined set of rules or signatures. This often involves manual feature engineering, where domain experts identify and define relevant features for detection. This process can be time-consuming and may not capture all relevant aspects of the data. Traditional Intrusion Detection Systems tend to struggle with adapting to new and evolving threats. Updates to rules or signatures are required to detect novel attack patterns, and these updates may lag emerging threats. Traditional Intrusion Detection Systems may face challenges in handling large volumes of data efficiently, especially when dealing with high-speed networks. Scaling traditional IDS can be resource intensive. Traditional Intrusion Detection Systems often require human intervention for rule or signature updates, as well as for fine-tuning the system based on new threats.

Above mentioned problems can be solved using DL models for Intrusion Detection Systems. DL models excel at detecting anomalies and can adapt to new and previously unseen attack patterns. These models automatically learn relevant features from raw data, removing the need for manual feature engineering. Exhibit adaptability by learning from data and adjusting their internal representations to new attack patterns. This makes them more effective in detecting previously unseen or zero-day attacks without frequent manual updates. DL models can be scalable and handle large datasets effectively, making them suitable for enterprise-level networks and high-speed traffic. They are well-suited for handling high-dimensional and complex data. DL models can operate with minimal human intervention once trained, as they can continuously learn and adapt to changes in the data distribution.

The rapid evolution and proliferation of cyber threats pose significant challenges to the security of computer networks. Intrusion Detection Systems (IDS) play a pivotal role in identifying and mitigating these threats by monitoring network activities for abnormal behaviours indicative of potential attacks. Traditional IDS techniques, such as signature-based detection and anomaly detection, have been fundamental to network security for a long time. Signature-based detection compares network traffic against known patterns of malicious activity, making it effective for detecting well-characterized attacks [1]. However, it struggles with zero-day attacks and variations of known threats. On the other hand, anomaly detection establishes a baseline of normal network behaviour and flags deviations as potential intrusions [2]. While effective in theory, anomaly detection often suffers from high false-positive rates and requires continuous updates to adapt to changing network conditions.

The limitations of traditional IDS methods have prompted the exploration of advanced techniques such as DL. DL models offer a paradigm shift by enabling automated feature extraction and learning from large-scale data, thereby enhancing detection accuracy and adaptability [3]. This paper aims to explore and evaluate the efficacy of DL models in IDS applications, discussing their potential to address the shortcomings of traditional approaches and contribute to the advancement of network security. DL, a subset of machine learning, utilizes neural networks with multiple layers to automatically learn features from data. The ability of DL models to handle large volumes of data and extract complex patterns makes them particularly suitable for intrusion detection.

This review aims to summarize the key DL models used for IDS, evaluate the performance of these models across different datasets, identify the strengths and weaknesses of each approach and suggest potential areas for future research.

2. Background and Evolution of Intrusion Detection

The concept of intrusion detection began to emerge in the 1970s as computer networks started to become more prevalent. Early systems focused on basic log analysis and auditing to detect unauthorized access.

In the 1980s, expert systems and rule-based systems started to be used for intrusion detection. Systems like the Haystack system, developed in the mid-1980s, used rule-based approaches to identify suspicious activities.

Pre-publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)

In the 1990s, signature-based detection emerged, involving the matching of known attack patterns against network traffic. Commercialization of intrusion detection systems began with the introduction of products like the Network Flight Recorder and RealSecure. In the late 1990s, anomaly-based detection gained attention as a complement to signature-based methods, focusing on deviations from normal behaviour. Hybrid approaches, combining signature and anomaly detection, started to emerge for more comprehensive threat detection.

In the early 2000s, Intrusion detection systems became a standard component of cybersecurity strategies for organizations. Open-source IDS solutions, such as Snort, gained popularity and contributed to the wider adoption of intrusion detection technology. In the mid-2000s, IDS systems began to integrate with Security Information and Event Management (SIEM) solutions for centralized log management and analysis. This integration improved the correlation of security events and enhanced the overall security posture.

The 2010s saw the increased use of machine learning techniques in intrusion detection systems. Behavioural analysis became more sophisticated, allowing IDS to adapt to evolving threats and identify previously unknown attack patterns.

Currently, modern intrusion detection systems continue to evolve, incorporating advanced threat detection capabilities. Cloud-based IDS solutions and threat intelligence feeds contribute to more robust and adaptive intrusion detection. Today, IDS plays a crucial role in the broader field of cybersecurity, helping organizations detect and respond to a wide range of cyber threats.

3. Deep Learning Models

Deep Learning (DL) has garnered significant attention across various domains for its ability to learn intricate patterns and representations directly from raw data. In the context of IDS, DL models have shown promising results in augmenting traditional detection techniques.

Convolutional Neural Networks (CNNs), initially designed for image processing tasks, have been successfully adapted to effectively analyse network traffic data. CNNs excel in capturing spatial dependencies within data, which makes them highly suitable for tasks like malware classification and anomaly detection in network traffic [4]. By applying filters to input data, CNNs can automatically extract meaningful features, reducing the reliance on manually crafted rules and signatures.

Awajan et al. presented a deep learning-based IDS specifically designed for IoT networks. This system addresses the unique challenges of IoT environments, such as resource constraints and diverse device types. The proposed model uses a CNN architecture to analyse IoT traffic patterns and detect anomalies with high precision. The study demonstrated the effectiveness of deep learning in securing IoT networks against various types of attacks [5].

Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, are another class of DL models that have demonstrated effectiveness in sequential data analysis. RNNs can capture time-based relationships in data sequences, which is key for detecting intrusion patterns that evolve [6]. The ability of LSTMs to retain information over extended time intervals enables them to detect subtle changes in network behaviour that may indicate ongoing or emerging security threats.

In addition to CNNs and RNNs, Generative Adversarial Networks (GANs) have emerged as an innovative method to enhance IDS capabilities. GANs are used to create artificial data that mimic real network traffic patterns, thus augmenting training datasets and improving the strength of IDS models in contradiction of adversarial attacks [7]. This approach leverages the power of DL to create diverse and realistic data samples, enabling IDS systems to generalize better to unseen threats and variations.

4. Deep Learning Models in Intrusion Detection

Wang et al. discussed the application of deep belief networks (DBNs) in network intrusion detection, demonstrating their capability to handle high dimensional data and detect complex attack patterns. DBNs, which are composed of multiple layers of restricted Boltzmann machines, are adept at capturing the hierarchical structure of network traffic data, thereby improving detection accuracy [4].

Pre-publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)

Yin et al. discovered the use of RNNs for anomaly detection in network traffic, emphasizing the ability of RNNs to model sequential data and capture temporal correlations between events. The study found that RNNs outperform traditional methods in scenarios where the order of network events plays a crucial role in identifying malicious behaviour [6].

Javaid et al. implemented a DL framework using autoencoders for intrusion detection, showing significant improvements in detection accuracy and reduction in false alarms compared to traditional machine learning techniques. Autoencoders, with their ability to learn compact representations of data, provide an efficient way to detect irregularities in large-scale network traffic [10].

Vinayakumar et al. evaluated various DL models for network intrusion detection, including CNNs, RNNs, and hybrid models, concluding that DL provides superior performance in identifying sophisticated attack vectors. The study highlighted the importance of model selection based on the specific features of the network environment and the nature of the threats [11].

Zhou et al. applied ensemble learning techniques with DL models to enhance intrusion detection accuracy, demonstrating the benefits of combining multiple models to capture diverse attack characteristics. By leveraging the strengths of different models, the ensemble approach achieves higher detection rates and robustness against various types of intrusions [12].

LeCun et al. provided a comprehensive summary of DL techniques and their applications, highlighting their potential in various domains by including network security. The review emphasized the flexibility and scalability of DL models, making them suitable for complex and dynamic network environments [13].
Roy et al. introduced a hybrid DL approach which combines CNNs and LSTMs for detecting DDoS attacks, showcasing the strengths of both architectures in processing spatial and temporal features of network traffic. The hybrid model effectively addresses the limitations of individual architectures, providing a robust solution for detecting distributed attacks [14].

Xiao et al. explored the usage of transfer learning to improve the generalization of intrusion detection systems across different network environments, addressing the challenge of data scarcity and variability. Transfer learning enables models trained on one dataset to be adapted to another., improving their performance in new and diverse settings [8].

Tang et al. explored the efficiency of DL models in detecting insider threats, emphasizing the importance of modelling user behaviour and access patterns to identify malicious activities from within the organization. The study demonstrated that DL models could effectively differentiate between normal and suspicious behaviour, even when the latter mimics legitimate activities [16].

Zhang et al. presented a framework that integrates feature selection and DL for intrusion detection, emphasizing the importance of selecting relevant features to enhance model performance and decrease computational complexity. By identifying the most important features, the proposed approach improves the efficiency and accuracy of the IDS [9].

Kim et al. proposed the use of LSTM RNNs for intrusion detection. Their research demonstrated that LSTMs could capture temporal dependencies in network traffic, making them particularly effective for detecting anomalies over time. This method achieved high accuracy in classifying normal and malicious activities, proving the effectiveness of RNNs in dynamic environments. The ability of LSTMs to remember long-term dependencies in the data makes them ideal for scenarios where the sequence of events is crucial for accurate detection [2].

Shone et al. introduced a novel DL approach to network intrusion detection, leveraging a combination of autoencoders and deep neural networks to enhance detection accuracy. The study demonstrated that DL models could effectively identify both known and unknown threats by learning complex patterns in network traffic data. Autoencoders reduce dimensionality and extract relevant features, which are then used by deep neural networks for classification. This approach proved particularly effective in identifying subtle anomalies that traditional methods might miss [1].

Siva Shankar et al. introduced a new optimization-based DL approach combined with artificial intelligence techniques to detect intrusion attacks in network systems. This method utilizes an optimization algorithm to

Pre-Publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)

enhance the efficacy of DL models, resulting in more accurate and efficient detection of network intrusions. The study highlighted the significant improvement in detection rates and reduction in false positives achieved through this approach. By integrating optimization techniques, the authors successfully fine-tune the parameters of the DL models, ensuring optimal performance in various network environments [15].

Lin et al. surveyed the application of GANs in network anomaly detection. GANs, consisting of a generator and a discriminator, are particularly useful for generating synthetic data that resembles real network traffic. This synthetic data can be used to train IDSs improving their ability to recognize novel attacks. The study highlighted the potential of GANs to enhance the robustness and generalization of intrusion detection models. By generating realistic attack scenarios, GANs help in training models that are better equipped to handle unseen threats [7].

Awajan et al. proposed a DL-based IDS specifically designed for IoT networks. This system addresses the unique challenges of IoT environments, such as resource constraints and diverse device types. The presented model uses a CNN architecture to analyse IoT traffic patterns and detect anomalies with high precision. The study demonstrated the effectiveness of DL in securing IoT networks against various types of attacks. The CNN-based model efficiently processes the high-dimensional data generated by IoT devices, identifying patterns indicative of malicious activities [5].

5. Datasets for Intrusion Detection

5.1 KDD Cup 1999 Dataset

The KDD Cup 1999 dataset is among the earliest and most extensively utilized datasets in research for Intrusion Detection Systems (IDS). It was derived from the DARPA 1998 dataset, which contains a variety of simulated network traffic, including both normal and attack traffic. This dataset has played an important role in the development and evaluation of IDS models over the years.
One of the primary strengths of the KDD Cup 1999 dataset is its extensive use in the literature. This widespread adoption makes it easier for researchers to compare their results with previous studies, providing a benchmark for evaluating new IDS models. Additionally, the dataset contains labelled data, which is crucial for supervised learning methods. Researchers can train their models on known attacks and normal traffic patterns, facilitating the development of accurate IDS systems.

However, the KDD Cup 1999 dataset has several notable weaknesses. Firstly, it is outdated and may not accurately represent modern network traffic. The nature of cyber-attacks has evolved significantly since the dataset was created, potentially limiting the effectiveness of models trained on this data when applied to contemporary network environments. Secondly, the dataset contains redundant records, which can bias the results and lead to overfitting.

This redundancy can inflate the performance metrics, giving a false sense of security regarding the model's accuracy. Finally, there are known issues with the quality and representativeness of the attacks in the dataset. Some attacks are either oversimplified or not representative of real-world scenarios, which can hinder the practical applicability of the IDS models developed using this data [17][18].

5.2 NSL-KDD Dataset

The NSL-KDD dataset was created as an upgraded version of the KDD Cup 1999 dataset, addressing many of its predecessor's shortcomings. It aims to provide a more accurate and effective benchmark for IDS research by reducing the issues of redundancy and class imbalance.

A significant strength of the NSL-KDD dataset is its reduced redundancy. Unlike the KDD Cup 1999 dataset, it eliminates duplicate records, which helps to mitigate the problem of biased results and overfitting. Additionally, the dataset has a more balanced distribution of classes, making it easier to train and evaluate models without the complications introduced by an uneven class distribution. This balance ensures that IDS models are more robust and perform better across different types of network traffic.

Despite these improvements, the NSL-KDD dataset still has limitations. It is based on the same outdated data as the KDD Cup 1999, which means it may not accurately reflect modern network traffic patterns and attack types. As cyber threats have evolved, relying solely on this dataset could limit the effectiveness of IDS models in contemporary settings. Researchers need to supplement this dataset with more current data to ensure their models

Pre-publication Copy – Not for Distribution

Copyright Notice (© 2025 hCAP Institute of Technology. All rights reserved.)

are relevant and effective against today's threats [19].

5. 3 UNSW-NB15 Dataset

The UNSW-NB15 dataset is a modern dataset designed to address the limitations of older datasets like KDD Cup 1999 and NSL-KDD. It was created using the IXIA PerfectStorm tool to generate realistic network traffic, including various attack scenarios, providing a more comprehensive resource for IDS research.

The primary strengths of the UNSW-NB15 dataset are its recency and comprehensiveness. It encompasses a diverse array of attack types, including both contemporary and sophisticated threats, making it more relevant for modern IDS development. The dataset includes both network flow and packet-based features, offering a rich feature set that enables detailed analysis and more accurate detection models. This comprehensive nature allows for better training and evaluation of IDS models in realistic network environments.

However, the UNSW-NB15 dataset is not without challenges. Its complexity can make it difficult for beginners to use effectively. The dataset requires significant preprocessing to be useful for machine learning models, including feature selection and data normalization. Additionally, the large size of the dataset can be computationally expensive to process, necessitating robust computing resources and efficient algorithms to handle the data effectively [20].

5.3 CICIDS2017 Dataset

The CICIDS2017 dataset was developed by the Canadian Institute for Cybersecurity and contains a mix of benign and malicious traffic, capturing a wide range of modern attack scenarios. It reflects contemporary network environments, making it a valuable resource for IDS research.

A notable strength of the CICIDS2017 dataset is its reflection of modern network traffic and attack patterns. This dataset includes detailed features, such as flow-based and packet-based characteristics, which enable the development of sophisticated and accurate IDS models. Researchers benefit from the comprehensive and up-to-date nature of the dataset, which is crucial for addressing current cyber threats.

However, the CICIDS2017 dataset also presents some challenges. Its large size can be computationally expensive to process, requiring significant storage and processing power. Additionally, the dataset requires a substantial effort for labelling and preprocessing to prepare it for machine learning applications. These steps are essential to ensure the dataset's effectiveness but can be time-consuming and resource-intensive [21].

5.4 CSE-CIC-IDS2018 Dataset

The CSE-CIC-IDS2018 dataset is a collaboration between the Communications Security Establishment (CSE) and the Canadian Institute for Cybersecurity. It contains diverse attack scenarios captured over multiple days, providing a comprehensive resource for IDS research.

One of the primary strengths of the CSE-CIC-IDS2018 dataset is its up-to-date nature. It includes modern attack techniques and realistic network traffic, making it highly relevant for current IDS development. The dataset captures both day-to-day network traffic and various attack scenarios, offering a rich and diverse set of data for training and evaluating IDS models.

However, the CSE-CIC-IDS2018 dataset also has its drawbacks. The significant amount of data can be challenging to handle, requiring robust computing resources for effective processing and analysis. Additionally, the dataset necessitates complex feature engineering to extract meaningful information from the raw data. This complexity can be a barrier for researchers, particularly those with limited experience in data preprocessing and feature extraction [22].

5.5 MAWI Dataset

The MAWI dataset is a collection of daily traffic captures from a trans-Pacific link between Japan and the United States. It is used for various network research purposes, offering a realistic view of network traffic over time.

A key strength of the MAWI dataset is its realistic nature. As it contains real-world traffic, it offers a true

Pre-publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)

representation of network scenarios, which is invaluable for developing IDS models that need to perform well in practical settings. The dataset is continuously updated, providing current data that reflects the latest trends in network traffic and attack patterns.

However, the MAWI dataset also poses significant challenges. One of the main issues is the lack of labeled data, which makes it difficult to use for supervised learning methods. Researchers must invest considerable effort into labeling and preprocessing the data, which can be demanding in terms of time and complexity. Also, the data can be noisy, requiring extensive cleaning and preprocessing to extract useful features for IDS model development [23].

Table 1 gives the comparison of different datasets.

Table 1. Comparison of different datasets

Dataset	Year	Type	Source	Traffic Type	Features	Size	Strengths	Weaknesses
KDD Cup 1999	1999	Simulated	DARPA 1998	Network traffic	41 features (e.g., protocol, service, flag, etc.)	4,898,431 instances	Widely used, extensive literature, labeled data	Outdated, redundant records, quality and representativeness issues
NSL-KDD	2009	Simulated	Improved KDD Cup 1999	Network traffic	41 features (same as KDD)	125,973 instances	Reduced redundancy, balanced class distribution	Still outdated, limited modern attack representation
UNSW-NB15	2015	Simulated	IXIA PerfectStorm	Network traffic	49 features (e.g., flow-based and packet-based)	2,540,044 instances	Recent, diverse attack types, rich feature set	Complex, significant preprocessing required
CICIDS 2017	2017	Simulated	Canadian Institute for Cybersecurity	Network traffic	80 features (e.g., flow-based and packet-based)	~3,000,000 instances	Modern traffic, detailed features	Large size, significant labeling and preprocessing effort
CSE-CIC-IDS2018	2018	Simulated	CSE & Canadian Institute for Cybersecurity	Network traffic	80+ features (e.g., time-based, protocol-based)	~16,000,000 instances	Up-to-date, diverse attack scenarios	Large volume, complex feature engineering required
MAWI	Ongoing	Real-world	MAWI Working Group	Real-world traffic	Depends on specific captures	Varies daily (large volume)	Realistic, continuously updated	Unlabeled, noisy, extensive preprocessing required

Pre-publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)

6 Challenges and Limitations

Despite their significant advancements, DL-based IDS face several challenges that must be addressed to facilitate their widespread adoption in practical settings. One of the primary concerns is the interpretability of DL models, as they often operate as black boxes for making it difficult to understand the rationale behind their decisions [9]. The lack of interpretability can hinder trust and acceptance among cybersecurity professionals and end-users, limiting the deployment of DL-based IDS systems in critical infrastructure and sensitive environments.

Furthermore, DL models require substantial computational resources and large volumes of labelled data for training, which may not always be immediately accessible in IDS applications [3]. This requirement poses scalability challenges, particularly in deploying IDS solutions across distributed or resource-constrained networks. Additionally, the security vulnerabilities associated with DL models, such as susceptibility to adversarial attacks, remain a significant concern [9]. Adversarial attacks can exploit vulnerabilities in DL architectures to manipulate or evade detection mechanisms, compromising the effectiveness of IDS systems in real-world scenarios.

7 Conclusions

DL models represent a significant advancement in the field of Intrusion Detection Systems, offering enhanced capabilities to detect and mitigate sophisticated cyber threats. While DL-based IDS have demonstrated promising results in various application domains, challenges related to interpretability, scalability, and security vulnerabilities must be resolved to facilitate their broader adoption in practical settings. Future research efforts should prioritize the development of robust, interpretable, and scalable DL-based IDS solutions that can effectively mitigate emerging cyber threats and safeguard network infrastructures.

References

[1] Shone N, Ngoc TN, Phai VD, Shi Q. A Deep Learning Approach to Network Intrusion Detection. IEEE Trans Emerg Top Comput Intell. 2018;2(1):41-50.
[2] Kim S, Kim H, Lee H. Long Short-Term Memory Recurrent Neural Network Classifier for Intrusion Detection. J Ambient Intell Human Comput. 2016;9(5):1293-1300.
[3] Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016.
[4] Wang W, Zhu M, Zeng X, Ye X, Sheng Y. Malware traffic classification using convolutional neural network for representation learning. 2017 Int Conf Inf Netw (ICOIN). 2017:712-717.
[5] Awajan A. A novel deep learning-based intrusion detection system for IOT networks. Computers. 2023;12(2):34.
[6] Yin C, Zhu Y, Fei J, He X. A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access. 2017;5:21954-21961.
[7] Lin W, Ye Z, Xu Y, He Z. Generative Adversarial Networks and Its Applications in Network Anomaly Detection: A Survey. J Netw Comput Appl. 2020;169:102767.
[8] Xiao Z, Zhang Z, Xu H. Enhancing intrusion detection using transfer learning: Model selection and performance evaluation. IEEE Access. 2019;7:141782-141791.
[9] Zhang J, Li Z, Guo X, Zhang X. Feature selection and parameter optimization for deep learning-based intrusion detection model. IEEE Access. 2019;7:75426-75436.
[10] Javaid A, Niyaz Q, Sun W, Alam M. A deep learning approach for network intrusion detection system. Proc 9th EAI Int Conf Bio-inspired Inf Commun Technol (BIONETICS). 2016:21-26.
[11] Vinayakumar R, Soman KP, Poornachandran P. Evaluating deep learning approaches to intrusion detection for cyber-security. Int Conf Adv Comput Commun Informatics (ICACCI). 2019:2291-2297.
[12] Zhou Y, Cheng G, Jiang S, Dai M. Building an efficient intrusion detection system based on feature selection and ensemble classifier. Comput Netw. 2020;174:107247.

Pre-publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)

[13] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
[14] Roy S, Cheung H, Ding Z. A hybrid deep learning approach for intrusion detection. IEEE Access. 2020;8:136431-136441.
[15] Siva Shankar S, Bui Thanh Hung, Prasun Chakrabarti, Tulika Chakrabarti, Gayatri Parasa. A novel optimization-based deep learning with artificial intelligence approach to detect intrusion attacks in network systems. Educ Inf Technol. 2024;29(4):3859-3883.
[16] Tang T, Chen J, Luo H. Insider threat detection based on deep learning. IEEE Access. 2018;6:11074-11083.
[17] Hettich S, Bay SD. The UCI KDD Archive. Univ Calif Dept Inf Comput Sci. 1999.
[18] Tavallaee M, Bagheri E, Lu W, Ghorbani AA. A Detailed Analysis of the KDD CUP 99 Data Set. Proc Second IEEE Symp Comput Intell Secur Def Appl (CISDA). 2009.
[19] Moustafa N, Slay J. UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems (UNSW-NB15 Network Data Set). Mil Commun Inf Syst Conf (MilCIS). 2015.
[20] Sharafaldin I, Lashkari AH, Ghorbani AA. Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. ICISSP. 2018.
[21] Lashkari AH, Draper Gil G, Mamun MSI, Ghorbani AA. Characterization of Tor Traffic Using Time Based Features. Proc 3rd Int Conf Inf Syst Secur Privacy (ICISSP). 2018.
[22] MAWI Working Group Traffic Archive. Samplepoint-G. 2006.

Author Details

Padmapani P. Tribhuvan, Ph.D.
Department of AI, hCAP Institute of Technology, Chhatrapati Sambhajinagar, India.

Research Interests: *Artificial Intelligence, Machine Learning, Sentiment Analysis*

Brief Bio: *Padmapani P. Tribhuvan is an AI Instructor in the Department of AI, hCAP Institute of Technology. She also worked as Associate Professor, DIEMS. She has 17.5 Years of Teaching experience. Her research focuses on artificial intelligence and machine learning applications.*

ORCID: 0000-0001-8437-0508

Amrapali P. Tribhuvan, Ph.D. Scholar
Department of Computer Science & IT,
Dr. Babasaheb Ambedkar Marathwada University, India

Research Interests: *Digital Heritage, Virtual Reality, Cultural Computing, Artificial Intelligence*

Brief Bio: *Amrapali P. Tribhuvan is a Ph.D. scholar specializing in the Digital Heritage of Ajanta Caves using Cultural Computing. She has 16 years of teaching experience in computer science. Her research focuses on integrating VR technology for digital heritage conservation and enhancing visitor experiences through AI-driven solutions. She has authored multiple research papers and is actively involved in projects related to immersive technology.*

ORCID: 0000-0001-5639-465X

Publication Copy – Not for Distribution

Copyright Notice (© *2025 hCAP Institute of Technology. All rights reserved.*)