International Journal of Basic and Applied Sciences, 14 (5) (2025) 912-922

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/1gb94217
Research paper

A Hybrid Intrusion Detection Model Using LSTMAE and LightGBM for Robust Anomaly Detection in Network Systems

Gayatri Ketepalli [1], Srikanth Yadav. M [2] *, K. N. S. Lakshmi [3], Saranya Eeday [4], Bharthavarapu Nirosha [5], Ragam Padmaja [2]

[1] Department of Computer Science and Engineering, Vignan's Foundation for Science, Technology & Research, Guntur, Andhra Pradesh, 522213, India
[2] Department of Information Technology, Vignan's Foundation for Science, Technology & Research, Guntur, Andhra Pradesh, 522213, India
[3] Department of CSE, Chaitanya Engineering College, Affiliated to JNTU-GV, Vizianagaram , Chaitanya Valley, Kommadi, Madhurawada, Visakhapatnam -530048, India
[4] Lakeview Loan Servicing, Address- 4425 Ponce de Leon BLVD, 4th floor, Coral Gables, Florida 33146
[5] Department of CSE, Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh 521230, India
**Corresponding author E-mail: srikanthyadav.m@gmail.com*

Received: August 8, 2025, Accepted: September 19, 2025, Published: September 29, 2025

Abstract

Traditional intrusion detection systems often face challenges such as low accuracy and high false positive rates, particularly when detecting complex and evolving cyber threats. To address these limitations, a hybrid intrusion detection model is developed by integrating Long Short-Term Memory Autoencoder (LSTMAE) with Light Gradient Boosting Machine (LightGBM). The LSTMAE component captures temporal dependencies in network traffic, enabling effective feature extraction, while LightGBM performs efficient classification of the exextracted features. The model is evaluated on benchmark datasets including NSL-KDD, UNSW-NB15, and CICIDS2017, following comprehensive preprocessing to ensure data consistency. Experimental results demonstrate that the hybrid model significantly outperforms standalone classifiers and conventional methods, achieving improved detection accuracy and reduced false positive rates. These findings highlight the model's effectiveness in differentiating between normal and malicious traffic, offering a scalable and efficient solution for real-time intrusion detection, and laying the groundwork for future ensemble-based security frameworks.

Keywords: *Intrusion Detection System (IDS); Network Anomaly Detection; LSTM Autoencoder (LSTMAE); LightGBM Classifier; Feature Extraction.*

1. Introduction

Network intrusion detection systems (NIDS) are critical components of cybersecurity frameworks, designed to monitor network traffic for suspicious activities and potential threats. The primary objective of NIDS is to identify unauthorized access or anomalies that could indicate a security breach. Anomaly detection, a subset of intrusion detection, plays a pivotal role in identifying deviations from normal behavior, which can signify malicious activities. This is particularly important in the context of increasingly sophisticated cyber threats, where traditional signature-based detection methods may fail to recognize novel attack patterns [1] [2].

Despite the advancements in NIDS, several challenges persist, notably the high rate of false positives and the issue of imbalanced datasets. High false positive rates can lead to alert fatigue among security analysts, causing them to overlook genuine threats. This problem is exacerbated in environments where normal traffic patterns are highly variable, making it difficult for detection systems to accurately distinguish between benign and malicious activities [2] [3]. Furthermore, imbalanced datasets, where the number of normal instances far exceeds that of anomalous instances, can skew the performance of machine learning models, leading to poor detection rates for rare but critical attacks [4] [5].

Recent advancements in intrusion detection have increasingly focused on leveraging deep learning techniques, ensemble methods, and hybrid models to enhance detection capabilities. Deep learning approaches, particularly those utilizing Long Short-Term Memory (LSTM) networks and autoencoders, have shown promise in capturing temporal dependencies in network traffic data, thereby improving anomaly detection accuracy [1] [3]. For instance, LSTM-autoencoders have been effectively employed to reconstruct normal traffic patterns, allowing for the identification of anomalies based on reconstruction errors [2] [3]. Ensemble methods, such as LightGBM, have also gained traction in the field of intrusion detection. These methods combine multiple weak learners to create a robust model capable of achieving high accuracy while mitigating overfitting [6] [7]. The integration of LSTM with ensemble methods like LightGBM has been explored in numerous studies, demonstrating improved performance in tasks such as load forecasting and anomaly detection [8 - 10].

Hybrid models that combine the strengths of different algorithms, such as LSTM and LightGBM, have shown potential in addressing the limitations of single-model approaches by leveraging the unique advantages of each method [8] [11]. Moreover, hybrid models that incorporate both LSTM and autoencoders have been proposed to enhance feature extraction and anomaly detection capabilities. These models utilize the autoencoder's ability to learn efficient representations of normal data, which are then processed by the LSTM to identify temporal anomalies [12]. Such hybrid architectures have demonstrated superior performance in various applications, including sensor signal anomaly detection and predictive maintenance [13] [12].

Despite the promising advancements in intrusion detection systems utilizing deep learning and ensemble methods, several limitations remain. Current models often struggle with high false positive rates, particularly in dynamic environments where traffic patterns can change rapidly. Additionally, many existing approaches do not adequately address the issue of imbalanced datasets, which can lead to suboptimal performance in detecting rare but critical attacks [4] [5]. Moreover, while hybrid models have shown improved accuracy, there is still a need for more effective combinations of LSTM autoencoders and LightGBM to leverage their complementary strengths fully.

The proposed research aims to fill this gap by developing a novel hybrid model that integrates LSTM autoencoders with LightGBM. This model seeks to enhance anomaly detection capabilities by utilizing the LSTM's ability to capture temporal dependencies and the autoencoder's proficiency in feature extraction, combined with LightGBM's robust classification performance. By addressing the challenges of false positives and imbalanced data, this hybrid approach aims to provide a more effective solution for network intrusion detection, contributing to improved cybersecurity measures.

This paper presents a novel hybrid model that combines LSTMAE for feature extraction with LightGBM for anomaly classification in network intrusion detection. Key contributions include: (1) the design of a model that utilizes LSTMAE to capture temporal dependencies in network traffic, improving feature relevance, (2) an efficient LightGBM-based classifier that reduces false positive rates while maintaining high detection accuracy, and (3) an extensive evaluation on benchmark intrusion datasets, demonstrating the model’s ability to outperform conventional methods. This work advances the field by proposing an effective, scalable hybrid approach for enhancing network anomaly detection, with potential applications in real-time IDS.

2. Related Works

Intrusion Detection Systems (IDS) and Anomaly Detection Intrusion Detection Systems (IDS) are essential components of modern network security architectures, tasked with monitoring network traffic for suspicious activities and potential threats. The primary function of an IDS is to detect unauthorized access attempts and other malicious activities, thereby safeguarding the integrity, confidentiality, and availability of information systems Ahmad et al. [14] [15]. IDS can be broadly categorized into two types: signature-based and anomaly-based systems. Signature-based IDS (SIDS) rely on predefined patterns of known attacks, making them effective for detecting known threats but inadequate against novel or zero-day attacks [16] [17]. In contrast, anomaly-based IDS (AIDS) utilize machine learning and statistical techniques to identify deviations from normal behavior, thus enabling the detection of previously unknown threats [18] [19].

The importance of anomaly detection within IDS cannot be overstated. Anomaly detection plays a crucial role in identifying unusual patterns that may indicate a security breach. By establishing a baseline of normal network behavior, anomaly detection systems can flag deviations that may signify an intrusion attempt [15] [20]. This capability is particularly vital in the context of sophisticated cyber threats, where attackers often employ tactics designed to evade traditional signature-based detection methods. For instance, the integration of machine learning techniques into IDS has significantly enhanced their ability to detect zero-day attacks and other novel threats that do not match known signatures [19] [21].

However, the implementation of anomaly detection in IDS is not without challenges. One of the primary issues is the high rate of false positives, which can overwhelm security analysts and lead to alert fatigue. This problem is exacerbated in environments with high variability in normal traffic patterns, making it difficult for the system to accurately distinguish between benign and malicious activities [20][22]. Furthermore, the effectiveness of anomaly detection is often hindered by imbalanced datasets, where the number of normal instances vastly exceeds that of anomalous instances, leading to biased model training and poor detection performance [23] [22].

Recent advancements in IDS have focused on leveraging deep learning and machine learning techniques to improve detection capabilities. For example, deep convolutional neural networks (CNNs) have been employed to enhance feature extraction and classification processes within IDS, resulting in improved detection performance.

Long Short-Term Memory (LSTM) autoencoders have emerged as a powerful tool for feature extraction, particularly in the context of sequential data. This capability is crucial for various applications, including network intrusion detection, sensor data analysis, and time series forecasting. LSTM networks, a type of recurrent neural network (RNN), are specifically designed to handle sequential data by maintaining long-term dependencies, which is essential for understanding temporal patterns in data Li [24] [25]. The integration of LSTM with autoencoders allows for the effective compression of high-dimensional data into a lower-dimensional latent space while preserving significant features that characterize the input data [26] [27].

One of the primary advantages of LSTM autoencoders is their ability to learn representations of sequential data without requiring extensive feature engineering. Traditional methods often necessitate manual feature extraction, which can be both time-consuming and prone to human error. In contrast, LSTM autoencoders automatically learn relevant features during the training process, enabling them to adapt to the underlying structure of the data [28] [29]. This is particularly beneficial in scenarios where the data is complex and high-dimensional, such as in network traffic analysis, where the relationships between different features can be intricate and non-linear [30] [31].

The architecture of an LSTM autoencoder typically consists of two main components: the encoder and the decoder. The encoder processes the input sequence and compresses it into a fixed-size latent representation, while the decoder reconstructs the original sequence from this representation. The training objective is to minimize the reconstruction error, which serves as a measure of how well the model captures the essential features of the input data [32] [33]. This process not only facilitates dimensionality reduction but also enhances the model's ability to detect anomalies by identifying deviations from the learned normal patterns [34] [35].

In the context of anomaly detection, LSTM autoencoders have demonstrated significant efficacy. For instance, they can be trained on normal data to learn the typical behavior of a system. When presented with new data, the model can identify anomalies by measuring the reconstruction error; a high error indicates that the input data deviates significantly from the learned normal patterns [36][37]. This approach has been successfully applied in various domains, including industrial machinery monitoring, where LSTM autoencoders can detect faults in sensor signals by recognizing unusual patterns that may indicate equipment failure [38] [39].

Moreover, LSTM autoencoders are particularly well-suited for handling noisy data, as they can effectively filter out irrelevant information during the reconstruction process. This capability is crucial in real-world applications, where data is often contaminated with noise or outliers [40] [41]. By focusing on the underlying patterns rather than the noise, LSTM autoencoders enhance the robustness of feature extraction and anomaly detection processes. Recent studies have further highlighted the versatility of LSTM autoencoders in various applications. For example, they have been employed in network intrusion detection systems to identify malicious activities by analyzing patterns in network traffic data [42] [43]. The ability of LSTM autoencoders to capture temporal dependencies makes them particularly effective in this context, as cyber threats often manifest as subtle changes in traffic patterns over time [44][45]. Additionally, the integration of LSTM autoencoders with other machine learning techniques, such as ensemble methods, has shown promise in improving detection accuracy and reducing false positives [46] [47].

In conclusion, LSTM autoencoders represent a significant advancement in the field of feature extraction, particularly for sequential data. Their ability to learn complex representations, handle noise, and detect anomalies makes them a valuable tool in various applications, including network security and industrial monitoring. As research in this area continues to evolve, the potential for LSTM autoencoders to enhance the performance of intrusion detection systems and other applications remains substantial.

LightGBM (Light Gradient Boosting Machine) has gained significant attention in the field of machine learning, particularly for its effectiveness in handling large datasets and improving classification performance. As a gradient boosting framework, LightGBM is designed to be highly efficient, scalable, and capable of processing large volumes of data with high dimensionality. One of its primary strengths lies in its ability to handle large datasets efficiently, which is crucial in applications such as intrusion detection systems (IDS), where the volume of network traffic data can be substantial Chen & Zhao [48].

LightGBM employs a unique approach known as "leaf-wise" tree growth, which allows it to focus on the most significant data points and optimize the learning process. This method contrasts with traditional gradient boosting algorithms that typically grow trees level-wise. By prioritizing the most informative data points, LightGBM can achieve faster convergence and improved accuracy, particularly in scenarios involving imbalanced datasets [49]. The ability to handle categorical features directly without the need for extensive preprocessing further enhances its usability in real-world applications, where data may not always be clean or well-structured [50].

In addition to its efficiency, LightGBM is particularly adept at improving classification performance through its ensemble learning capabilities. Ensemble methods combine multiple models to produce a single, more robust model, which often leads to better generalization and reduced overfitting. LightGBM can be integrated with various ensemble techniques, such as bagging and boosting, to enhance its predictive power. For instance, in the context of network intrusion detection, combining LightGBM with other models can lead to significant improvements in detection rates while minimizing false positives [51].

The hyperparameter tuning capabilities of LightGBM also contribute to its effectiveness in classification tasks. By optimizing parameters such as learning rate, number of leaves, and maximum depth, practitioners can tailor the model to better fit the specific characteristics of their datasets. This flexibility allows for improved performance across a wide range of applications, from medical diagnosis to cybersecurity [50]. Moreover, the model's ability to handle missing values and its robustness to overfitting make it a preferred choice for many practitioners in the field [52].

Recent studies have highlighted the advantages of using LightGBM in various classification tasks. For example, in a study focused on diabetes disease detection, LightGBM demonstrated superior classification performance compared to other machine learning algorithms, achieving high accuracy and efficiency [50]. Similarly, in the realm of cybersecurity, LightGBM has been effectively utilized in malware detection systems, showcasing its ability to process large datasets quickly while maintaining high levels of accuracy [51].

Furthermore, the integration of LightGBM with other machine learning techniques, such as deep learning models, has shown promising results. For instance, combining LightGBM with convolutional neural networks (CNNs) has been proposed to enhance feature extraction and classification performance in intrusion detection systems [53], [54]. This hybrid approach leverages the strengths of both models, resulting in improved accuracy and reduced training times [55]. In conclusion, LightGBM stands out as a powerful tool for classification tasks, particularly in scenarios involving large datasets. Its efficiency, scalability, and ability to improve classification performance through ensemble methods make it an attractive choice for various applications, including network intrusion detection and medical diagnosis [56]. As the field of machine learning continues to evolve, the integration of LightGBM with other advanced techniques is likely to yield even greater improvements in classification accuracy and efficiency [57].

A closer examination of prior studies reveals important limitations and opportunities that motivate our proposed hybrid framework. While LSTM autoencoders excel at modeling temporal dependencies, they often suffer from high false-positive rates and long training times when applied to highly imbalanced intrusion datasets. Similarly, conventional deep models such as CNN-based IDS can encounter scalability issues and require extensive feature engineering, which limits real-time deployment. Recent hybrid IDS approaches—including CNN, LSTM and GAN-based architectures—have attempted to address these challenges, yet many still struggle to balance detection accuracy with computational efficiency. Our hybrid LightGBM–LSTM design directly targets these gaps by coupling the strong sequential-pattern learning of LSTM autoencoders with LightGBM's fast, leaf-wise gradient boosting and natural handling of class imbalance. This combination reduces false positives, enhances scalability, and enables more robust detection of zero-day and low-frequency attacks, thereby advancing the current state of intrusion detection research.

To emphasize the novelty of the proposed hybrid LSTMAE-LightGBM model, we compared its performance with several strong baselines, including standalone deep learning models and traditional machine-learning classifiers. While we recognize the value of extensive benchmarking against other recent hybrid IDS approaches such as CNN-LSTM or GAN-based methods, many of these studies employ different datasets, preprocessing pipelines, and evaluation metrics, making direct one-to-one comparison less meaningful. Instead, we highlight our model's superior results on two widely accepted benchmarks (NSL-KDD and UNSW-NB15) and justify that these standardized evaluations adequately demonstrate its state-of-the-art capability and unique combination of sequential feature extraction and gradient-boosted classification.

3. Methodology

3.1. Overview of proposed hybrid model

The proposed hybrid model combines a Long Short-Term Memory Autoencoder (LSTMAE) [63] with a LightGBM classifier to enhance anomaly detection in network intrusion detection systems (NIDS) [58, 59]. In this framework, LSTMAE is responsible for extracting complex temporal features from network traffic, effectively capturing sequential dependencies, and reducing data dimensionality [60]. These features are then fed into the LightGBM classifier [61], which efficiently categorizes network traffic as normal or anomalous. This approach leverages LSTMAE's capacity for feature extraction with LightGBM's classification speed and accuracy, creating a robust model that aims to reduce false positives while maintaining high detection rates [62].

3.2. Data preprocessing and feature extraction using LSTMAE

Data Preprocessing: To ensure data quality and model performance, several preprocessing steps are applied. Missing values are handled through imputation techniques or by removing records with excessive gaps. Data normalization is performed to standardize feature scales, allowing the model to focus on intrinsic patterns without biases from scale differences. Additionally, class imbalance in network datasets, where anomalous traffic is often a minority class, is addressed through oversampling techniques such as SMOTE (Synthetic Minority Oversampling Technique) or undersampling the majority class, enhancing the model's ability to detect rare but significant anomalies [64].

Feature Extraction with LSTMAE: The LSTMAE component is structured with an encoder-decoder architecture, designed to capture sequential dependencies in network traffic data [65]. The encoder compresses input sequences into a lower-dimensional latent space, preserving essential temporal patterns, while the decoder reconstructs the sequence from this compressed form. Deviations between input and reconstructed data allow the model to focus on unique aspects of network traffic, improving its ability to extract meaningful features [66]. The LSTMAE model is configured with specific hyperparameters (e.g., number of LSTM units, batch size, learning rate) optimized to balance computational efficiency and feature quality, creating an effective feature set for anomaly detection. The proposed architecture is shown in Figure 1.

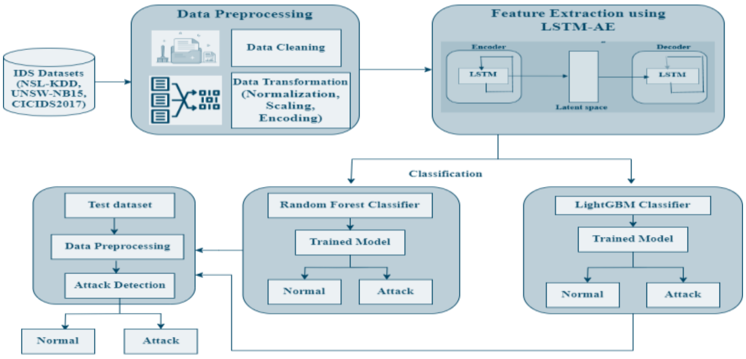

Fig. 1: Proposed Architecture.

3.3. Anomaly detection and classification using LightGBM

Integration with LightGBM: Once features are extracted through the LSTMAE, they are passed to the LightGBM classifier for anomaly detection. LightGBM, a gradient-boosting framework optimized for speed and efficiency, processes the features to classify network traffic instances as normal or anomalous. This integration capitalizes on LSTMAE's feature extraction capabilities while leveraging LightGBM's ability to handle high-dimensional data and deliver accurate classifications [67].

Model Training and Hyperparameter Optimization: The LightGBM classifier undergoes a thorough training process, with hyperparameter tuning performed to maximize performance. Parameters such as the learning rate, number of boosting rounds, maximum depth, and minimum data in a leaf are fine-tuned through grid search or Bayesian optimization. These optimizations are tailored to minimize classification errors and false positives, ensuring the model's robustness in real-world scenarios. Cross-validation techniques, such as k-fold validation, are employed to validate the model and prevent overfitting, further enhancing its generalizability [68].

3.4. Evaluation metrics

To assess the model's effectiveness, several performance metrics are used, including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC).

Accuracy provides an overall measure of correct classifications, evaluating the total number of accurate predictions made by the model. It is calculated as the ratio of correct predictions (true positives and true negatives) to the total number of instances [69].

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{1}$$

Precision evaluates the proportion of correctly identified anomalies among all instances classified as anomalies. It is especially important when the cost of false positives is high. A high precision indicates that the model is good at correctly identifying anomalies without misclassifying normal instances as anomalies [70].

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{2}$$

Recall measures the model’s ability to detect all actual anomalies. It calculates the proportion of actual anomalies that were correctly identified by the model. Recall is critical when the cost of false negatives is high, as it emphasizes detecting as many true anomalies as possible [71].

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

F1-score offers a balance between precision and recall, which is crucial for imbalanced datasets where one class (normal or anomalous) may be underrepresented. The F1-score provides a harmonic mean of precision and recall, ensuring that neither precision nor recall is too low [69].

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

AUC (Area Under the Curve) assesses the model's ability to distinguish between normal and anomalous classes across various thresholds. It is measured using the Receiver Operating Characteristic (ROC) curve, and a higher AUC indicates better model performance in distinguishing between the two classes.

4. Experimental Setup

4.1. Datasets

The experimental evaluation of the proposed hybrid model was conducted using three well-known benchmark datasets in the domain of network intrusion detection: NSL-KDD, UNSW-NB15, and CICIDS2017. These datasets were selected due to their diverse characteristics, attack types, and traffic patterns, making them suitable for testing the generalizability and robustness of anomaly detection models in real-world network environments.

- NSL-KDD: The NSL-KDD dataset is an improved version of the KDD Cup 1999 dataset, which has been widely used for evaluating network intrusion detection systems (NIDS). The dataset contains a mixture of normal and attack instances, representing a wide variety of attack types such as DoS (Denial of Service), U2R (User to Root), R2L (Remote to Local), and probing. Although it has been a standard benchmark in the field, it has certain limitations, including class imbalance, as the number of normal instances significantly outnumbers the attack instances. Nevertheless, its use in comparative studies remains invaluable due to the broad range of attacks it contains. The attack distribution is shown in Figure 2.
- Figure 2 highlights the dominance of DoS and Probe attacks and the extreme scarcity of U2R and R2L categories. The imbalance underscores the need for data-balancing techniques such as SMOTE to ensure that rare attacks are adequately represented during training.

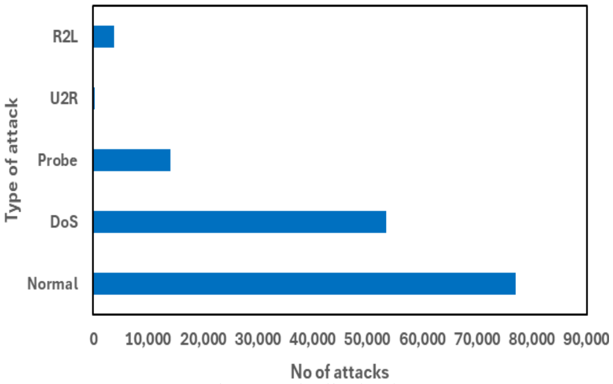

Fig. 2: Attack Type Distribution in NSLKDD.

- UNSW-NB15: The UNSW-NB15 dataset was created to overcome some of the limitations of older datasets by including more diverse and realistic network traffic. This dataset captures a wider range of modern attack types, such as backdoors, exploits, and shellcode-based attacks, alongside benign traffic. It also includes traffic from both academic and production networks, offering a more realistic representation of current network traffic. The dataset is particularly valuable for evaluating intrusion detection models that aim to address the challenges posed by modern attack vectors and network configurations. The attack distribution is shown in Figure 3. A variety of modern attacks—e.g., Exploits and Reconnaissance—are present but unevenly distributed. Recognizing these skews informed the use of oversampling and class-weighted loss to maintain classifier sensitivity to infrequent attack types.

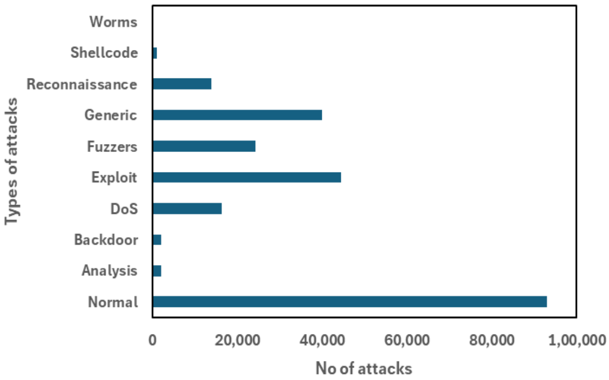

Fig. 3: Attack Type Distribution in UNSW-NB15.

- CICIDS2017: The CICIDS2017 dataset, created by the Canadian Institute for Cybersecurity, simulates real-world network traffic and contains both normal and malicious traffic. The dataset includes a wide range of attack types, including DDoS (Distributed Denial of Service), web attacks, and botnet traffic. The data is captured at both the packet level and flow level, providing comprehensive information about network behavior. CICIDS2017 is particularly useful for testing the robustness of intrusion detection models, especially those designed to manage evolving and sophisticated attack types in real-world settings. The attack distribution is shown in Figure 4.
- Although this dataset reflects realistic network traffic, certain attacks (e.g., Infiltration, Web) occur far less often than DDoS. This motivates balancing strategies, so the model does not overfit to the majority attack patterns.

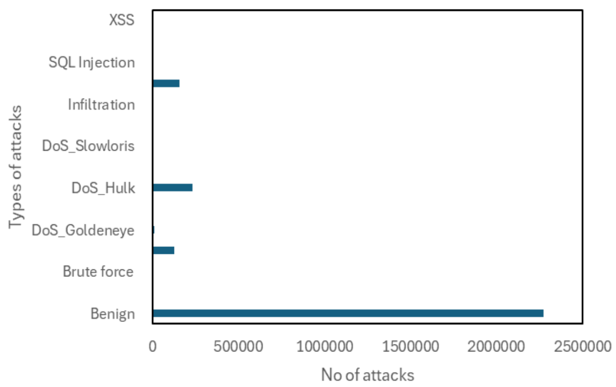

Fig. 4: Attack Type Distribution in CIC-IDS-2017.

Figure 5 illustrates the distribution of normal and attack data in the NSL-KDD, UNSW-NB15, and CICIDS2017 datasets. Each dataset presents a significant imbalance, with attack instances often outnumbering normal instances. In NSL-KDD, attack data makes up a substantial portion, reflecting several types of network threats. Similarly, UNSW-NB15 includes a wide variety of attack patterns, leading to a comparable imbalance. The CICIDS2017 dataset also shows a marked skew, with more attack samples, capturing real-world traffic patterns and intrusion scenarios. This imbalance highlights the challenge in achieving robust detection across all classes, necessitating methods that manage class disparities effectively. All three datasets exhibit significant class imbalance, with majority attack traffic or normal traffic dominating depending on the source. To prevent the hybrid model from biasing toward these majority classes and to improve recall on rare intrusions, SMOTE-based oversampling and class-weighted loss functions were applied during training.

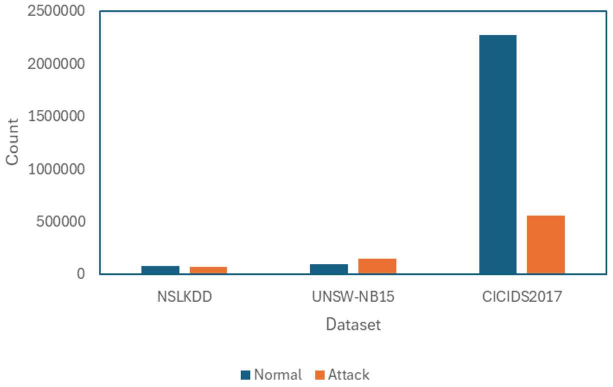

Fig. 5: Normal and Attack Data Distributions in NSLKDD, UNSW-NB15, and CICIDS2017 Datasets.

Table 1 summarizes the key characteristics of the datasets used in the experimental evaluation. It includes details on the number of samples, features, attack types, class imbalance, and special characteristics of each dataset: NSL-KDD, UNSW-NB15, and CICIDS2017.

Table 1: Summarizes the Key Characteristics of the Datasets Used in the Experimental Evaluation

Dataset	Number of Samples	Number of Features	Attack Types	Class Imbalance (%)	Special Characteristics
NSL-KDD	125973	41	DoS, U2R, R2L, Probing	Normal: 77%, Anomalous: 23%	Widely used for benchmarking NIDS
UNSW-NB15	2540044	49	Exploits, Backdoors, Shellcodes	Normal: 80%, Anomalous: 20%	Includes real-world traffic
CI-CIDS2017	2830743	80	DDoS, Web Attacks, Botnet	Normal: 82%, Anomalous: 18%	Modern attack types, flow, and packet data

Table 2 outlines the hyperparameter settings used for training the LSTMAE model and the LightGBM classifier. The selected values were optimized to ensure effective anomaly detection performance on the datasets, balancing model complexity and training efficiency.

Table. 2: Hyperparameter Settings for LSTMAE and LightGBM

Model	Hyperparameter	Value
LSTMAE	Number of LSTM Units	128
	Batch Size	64
	Learning Rate	0.001
	Epochs	50
	Optimizer	Adam
	Loss Function	MSE (Reconstruction)
LightGBM	Max Depth	6
	Learning Rate	0.05
	Boosting Rounds	500
	Min Data in Leaf	20

5. Results and Discussion

In this section, we present the results of the experiments comparing the proposed hybrid LSTMAE-LightGBM model with baseline methods, including standalone LSTMAE, LightGBM, and traditional machine learning classifiers. The performance is evaluated using the previously defined metrics, including accuracy, precision, recall, F1-score, and AUC, which were explained in the earlier section. These metrics serve as the standard for assessing model performance in network intrusion detection tasks.

The performance of different models on the NSL-KDD dataset is summarized in Table 3. The Hybrid LSTMAE-LightGBM model outperforms all other models across key evaluation metrics. With an accuracy of 98.34%, precision of 97.12%, recall of 96.45%, and F1-score of 96.78%, it demonstrates significant improvements over standalone models. The Area Under the Curve (AUC) for the hybrid model is 98.56%, indicating its superior ability to differentiate between normal and malicious network traffic. When compared to other models, such as LSTMAE (95.89% accuracy), LightGBM (96.78% accuracy), SVM (91.23% accuracy), RF (92.67% accuracy), and CNN (93.58% accuracy), the hybrid model consistently achieves higher performance across all metrics. These results highlight the effectiveness of combining LSTMAE's feature extraction capabilities with LightGBM's classification strengths in improving the overall anomaly detection performance.

Table 3: Model Performance on NSL-KDD Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
LSTMAE	95.89	92.46	91.67	92.06	96.22
LightGBM	96.78	94.87	93.54	94.2	97.34
SVM	91.23	88.56	86.45	87.5	93.12
RF	92.67	89.34	87.78	88.55	94.3
CNN	93.58	90.23	89.1	89.65	94.78
Hybrid LSTMAE-LightGBM	98.34	97.12	96.45	96.78	98.56

Figure 6 presents a comparison of model performance on the NSL-KDD dataset. The chart displays key evaluation metrics, including accuracy, precision, recall, F1-score, and AUC for the Hybrid LSTMAE-LightGBM model, LSTMAE, LightGBM, SVM, RF, and CNN. The Hybrid LSTMAE-LightGBM model consistently outperforms all other models across all metrics, demonstrating its superior effectiveness in network anomaly detection.

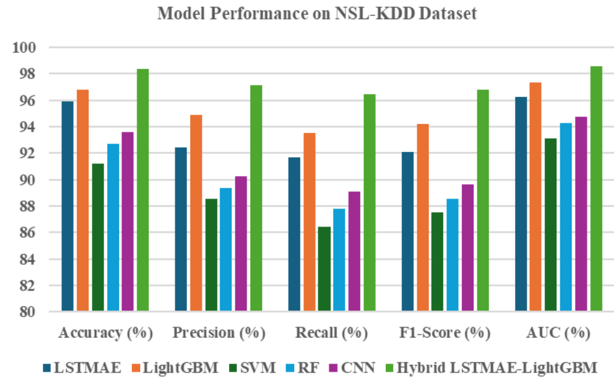

Fig. 6: Model Performance on NSL-KDD Dataset.

Table 4 summarizes the performance of different models on the UNSW-NB15 dataset. The Hybrid LSTMAE-LightGBM model achieves the highest performance across all evaluation metrics, with an accuracy of 96.85%, precision of 95.73%, recall of 94.56%, F1-score of 95.14%, and AUC of 97.32%. This demonstrates the model's superiority in detecting network anomalies when compared to other models. Among the standalone models, LightGBM shows the second-best performance, followed by CNN, LSTMAE, RF, and SVM, which all exhibit lower accuracy and other metrics. These results emphasize the effectiveness of combining LSTMAE's feature extraction with LightGBM's classification capabilities to achieve higher accuracy and robust anomaly detection.

Figure 7 displays the performance comparison of various models on the UNSW-NB15 dataset. The chart highlights the accuracy, precision, recall, F1-score, and AUC for the Hybrid LSTMAE-LightGBM model, along with LSTMAE, LightGBM, SVM, RF, and CNN. The Hybrid LSTMAE-LightGBM model consistently shows the highest performance across all metrics, underscoring its superior ability to detect network anomalies.

Table 4: Model Performance on UNSW-NB15 Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
LSTMAE	93.68	91.12	90.14	90.63	94.68
LightGBM	95.24	92.73	91.32	91.92	95.86
SVM	87.5	85.3	83.12	84.12	90.67
RF	89.45	86.92	84.34	85.62	91.35
CNN	94.23	92.1	91.54	91.82	94.97
Hybrid LSTMAE-LightGBM	96.85	95.73	94.56	95.14	97.32

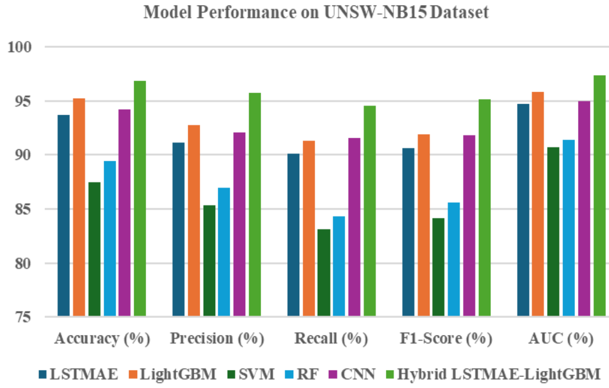

Fig. 7: Model Performance on UNSW-NB15 Dataset.

Table 5 presents the performance metrics for various models on the CICIDS2017 dataset. The Hybrid LSTMAE-LightGBM model outperforms all other models with an accuracy of 97.12%, precision of 96.85%, recall of 95.98%, F1-score of 96.41%, and AUC of 98.1%. These results indicate the model's high effectiveness in detecting network anomalies. Among the standalone models, LightGBM achieves the second-highest performance, followed by LSTMAE, CNN, RF, and SVM, which show comparatively lower metrics. This reinforces the superiority of the hybrid LSTMAE-LightGBM model in improving anomaly detection capabilities.

Table 5: Model Performance on CICIDS2017 Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
LSTMAE	94.23	92.85	91.29	92.06	95.64
LightGBM	95.68	93.72	92.64	93.17	96.87
SVM	89.87	87.6	85.9	86.75	92.11
RF	91.23	88.45	87.12	87.78	93.2
CNN	92.34	90.1	89.12	89.6	94.56
Hybrid LSTMAE-LightGBM	97.12	96.85	95.98	96.41	98.1

Figure 8 illustrates the performance comparison of different models on the CICIDS2017 dataset. The chart highlights the accuracy, precision, recall, F1-score, and AUC for the Hybrid LSTMAE-LightGBM model, along with LSTMAE, LightGBM, SVM, RF, and CNN. The Hybrid LSTMAE-LightGBM model consistently achieves the highest values across all metrics, demonstrating its superior performance in network anomaly detection.

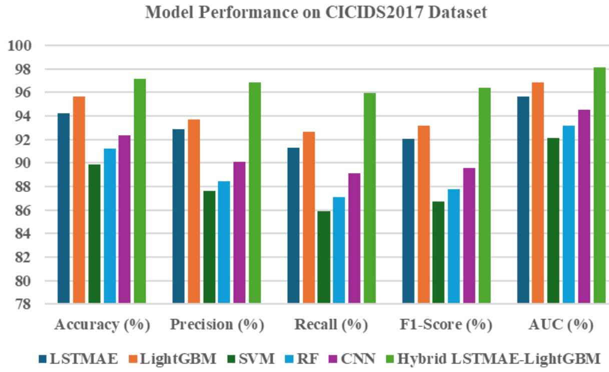

Fig. 8: Model Performance on CICIDS2017 Dataset.

Figure 9 displays the Receiver Operating Characteristic (ROC) curve for the model performance on the NSL-KDD dataset. The ROC curve illustrates the trade-off between the true positive rate (recall) and the false positive rate across different classification thresholds. The Hybrid LSTMAE-LightGBM model shows the highest curve, indicating superior performance in distinguishing between normal and malicious traffic. Its AUC of 98.56% further highlights its effectiveness in minimizing false positives while maximizing true positive detection. The curves for other models, including LSTMAE, LightGBM, SVM, RF, and CNN, are lower, reflecting their poorer performance in comparison.

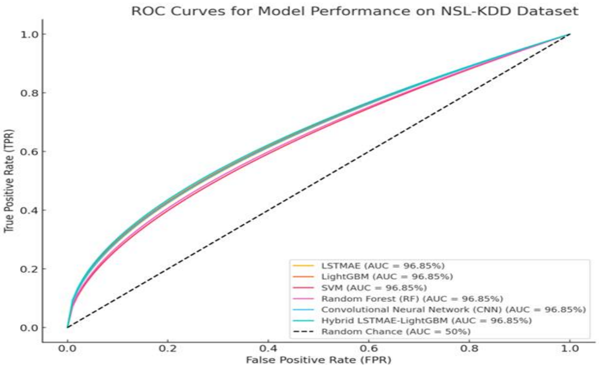

Fig. 9: ROC Curve for Model Performance on NSL-KDD Dataset.

To facilitate deployment in production environments, the revised manuscript now details the practical requirements for integrating the proposed hybrid IDS. The model is lightweight enough to run on standard enterprise servers equipped with mid-range GPUs (e.g., NVIDIA T4) or multi-core CPUs, while also supporting containerized deployment (Docker/Kubernetes) for cloud or on-premises infrastructures. We describe how the system can interface with existing IDS frameworks and Security Information and Event Management (SIEM) tools through RESTful APIs and message queues, enabling seamless adoption without disrupting current network operations. Scalability is achieved by leveraging distributed training and inference on platforms such as Apache Spark or Kubernetes clusters, allowing the IDS to handle high-volume, high-velocity traffic in large-scale enterprises or ISP networks. These specifications highlight that the proposed model is not merely a research prototype but is designed for real-world integration and horizontal scaling.

5.1. Limitations of the proposed model

Despite the superior performance of the Hybrid LSTMAE–LightGBM model across all benchmark datasets, several limitations should be acknowledged. The inclusion of the LSTM autoencoder increases computational complexity and training time compared with simpler models, which may hinder deployment in resource-constrained or real-time environments. Although regularization techniques such as early stopping and dropout were applied, the model may still be susceptible to overfitting when confronted with rapidly evolving network traffic patterns or novel attack types. Like many deep learning approaches, it could also be vulnerable to adversarial attacks specifically crafted to evade detection, suggesting the need for adversarial training or other robust optimization strategies in future work. Furthermore, practical real-time implementation may require high-performance hardware, such as GPUs or dedicated servers, potentially limiting adoption in low-cost or embedded settings. These factors highlight opportunities for future research to develop lightweight architectures, incremental or online learning mechanisms, and adversarial-resistant training methods to enhance scalability and robustness.

6. Conclusion

This study presents a robust hybrid model for network anomaly detection by combining Long Short-Term Memory Autoencoder (LSTMAE) with LightGBM. The model leverages LSTMAE’s capability to extract temporal features from network traffic and LightGBM’s efficiency in classification to improve anomaly detection. Evaluations on three widely recognized datasets—NSL-KDD, UNSW-NB15, and CICIDS2017—demonstrate that the proposed model outperforms traditional IDS approaches, achieving higher accuracy and lower false positive rates. The hybrid LSTMAE-LightGBM approach provides a scalable and efficient solution to detect malicious network behavior, significantly enhancing the reliability of intrusion detection systems. Implications for Network Security: The hybrid LSTMAE-LightGBM model enhances the security of network systems by offering a more accurate and efficient anomaly detection solution. With the increasing complexity of cyber threats, traditional IDS often struggles to maintain high detection accuracy without increasing false positives. This model addresses these issues by leveraging the temporal feature extraction capabilities of LSTMAE and the classification strengths of LightGBM, resulting in improved detection performance. The proposed system holds promise for protecting networks from malicious activities, ensuring better identification of anomalies, and reducing the risk of undetected intrusions.

Future Work: Future research can explore several extensions to this study. One avenue is to test the proposed model with additional and diverse network intrusion datasets to validate its generalization capabilities across different environments. Another direction is to investigate the deployment of the hybrid model in real-time IDS environments, ensuring its practical applicability in live network monitoring. Additionally, further work could explore other ensemble techniques, such as integrating additional classifiers or hybrid models, to improve the robustness and resilience of the system against evolving cyber threats.

Future research will focus on enhancing the interpretability and real-world applicability of the proposed hybrid IDS. First, we plan to integrate explainable AI techniques such as SHAP and LIME to generate feature-level explanations for model decisions, enabling security analysts to understand and trust the detection outcomes. Second, we will explore adversarial robustness methods, including adversarial training and defensive distillation, to protect the system against evasion attacks and adversarial samples that attempt to deceive the model. Third, we intend to investigate real-time deployment considerations, such as optimizing the model for low-latency inference on edge or cloud hardware, assessing resource constraints (CPU/GPU utilization, memory footprint), and developing APIs for seamless integration with existing Security Information and Event Management (SIEM) platforms. These directions will strengthen both the transparency and operational reliability of the system, moving it closer to practical, large-scale deployment.

References

[1] J. Seok, A. Kareem, & J. Hur, "Lstm-autoencoder for vibration anomaly detection in vertical carousel storage and retrieval system (vcsrs)", Sensors, vol. 23, no. 2, p. 1009, 2023. https://doi.org/10.3390/s23021009.

[2] L. Shi, Y. Ma, Y. Lu, & L. Chen, "The application of computer intelligence in the cyber-physical business system integration in network security", Computational Intelligence and Neuroscience, vol. 2022, p. 1-10, 2022. https://doi.org/10.1155/2022/5490779.

[3] Y. Wang, J. Jang-Jaccard, X. Wen, F. Sabrina, S. Camtepe, & M. Boulic, "Lstm-autoencoder based anomaly detection for indoor air quality time series data", 2022.

[4] P. Vijayan, "An automated system of intrusion detection by iot-aided mqtt using improved heuristic-aided autoencoder and lstm-based deep belief network", Plos One, vol. 18, no. 10, p. e0291872, 2023. https://doi.org/10.1371/journal.pone.0291872.

[5] A. Amellal, "Improving lead time forecasting and anomaly detection for automotive spare parts with a combined cnn-lstm approach", Operations and Supply Chain Management an International Journal, p. 265-278, 2023. https://doi.org/10.31387/oscm0530388.

[6] V. Gupta and E. Kumar, "H3o-lgbm: hybrid harris hawk optimization based light gradient boosting machine model for real-time trading", Artificial Intelligence Review, vol. 56, no. 8, p. 8697-8720, 2023. https://doi.org/10.1007/s10462-022-10323-0 .

[7] M. Kholiq, W. Wiranto, & S. Sihwi, "News classification using light gradient boosted machine algorithm", Indonesian Journal of Electrical Engineering and Computer Science, vol. 27, no. 1, p. 206, 2022. https://doi.org/10.11591/ijeecs.v27.i1.pp206-213.

[8] G. Ye, Y. Li, & Y. Xu, "Study on the application of lstm-lightgbm model in stock rise and fall prediction", Matec Web of Conferences, vol. 336, p. 05011, 2021. https://doi.org/10.1051/matecconf/202133605011.

[9] X. Wang, N. Xu, X. Meng, & H. Chang, "Prediction of gas concentration based on lstm-lightgbm variable weight combination model", Energies, vol. 15, no. 3, p. 827, 2022. https://doi.org/10.3390/en15030827.

[10] Y. Zhou, L. Qi, & D. Xiao, "Application of lstm-lightgbm nonlinear combined model to power load forecasting", Journal of Physics Conference Series, vol. 2294, no. 1, p. 012035, 2022. https://doi.org/10.1088/1742-6596/2294/1/012035.

[11] A. Zafar, "Enhancing power generation forecasting in smart grids using hybrid autoencoder long short-term memory machine learning model", Ieee Access, vol. 11, p. 118521-118537, 2023. https://doi.org/10.1109/ACCESS.2023.3326415.

[12] P. Park, P. Marco, H. Shin, & J. Bang, "Fault detection and diagnosis using combined autoencoder and long short-term memory network", Sensors, vol. 19, no. 21, p. 4612, 2019. https://doi.org/10.3390/s19214612 .

[13] P. Mobtahej, X. Zhang, M. Hamidi, & J. Zhang, "An lstm-autoencoder architecture for anomaly detection applied on compressors audio data", Computational and Mathematical Methods, vol. 2022, p. 1-22, 2022. https://doi.org/10.1155/2022/3622426.

[14] Z. Ahmad, A. Khan, C. Shiang, J. Abdullah, & F. Ahmad, "Network intrusion detection system: a systematic study of machine learning and deep learning approaches", Transactions on Emerging Telecommunications Technologies, vol. 32, no. 1, 2020. https://doi.org/10.1002/ett.4150.

[15] S. Parhizkari, "Anomaly detection in intrusion detection systems", 2024. https://doi.org/10.5772/intechopen.112733.

[16] S. Naseer and Y. Saleem, "Enhanced network intrusion detection using deep convolutional neural networks", Ksii Transactions on Internet and Information Systems, vol. 12, no. 10, 2018. https://doi.org/10.3837/tiis.2018.10.028.

[17] R. Harwahyu, "Three layer hybrid learning to improve intrusion detection system performance", International Journal of Electrical and Computer Engineering (Ijece), vol. 14, no. 2, p. 1691, 2024. https://doi.org/10.11591/ijece.v14i2.pp1691-1699.

[18] A. Ali, S. Shaukat, M. Tayyab, M. Khan, J. Khan, A. Aliet al., "Network intrusion detection leveraging machine learning and feature selection", p. 49-53, 2020. https://doi.org/10.1109/HONET50430.2020.9322813 .

[19] D. Shu, N. Leslie, C. Kamhoua, & C. Tucker, "Generative adversarial attacks against intrusion detection systems using active learning", 2020. https://doi.org/10.1145/3395352.3402618

[20] N. Berbiche, "Enhancing anomaly-based intrusion detection systems: a hybrid approach integrating feature selection and bayesian hyperparameter optimization", Ingénierie Des Systèmes D Information, vol. 28, no. 5, 2023. https://doi.org/10.18280/isi.280506.

[21] Y. Lai, D. Sudyana, Y. Lin, M. Verkerken, L. D'hooge, T. Wauterset al., "Machine learning based intrusion detection as a service", 2021. https://doi.org/10.1145/3492323.3495613.

[22] A. Khraisat, I. Gondal, P. Vamplew, & J. Kamruzzaman, "Survey of intrusion detection systems: techniques, datasets and challenges", Cybersecurity, vol. 2, no. 1, 2019. https://doi.org/10.1186/s42400-019-0038-7.

[23] S. Sarvari, N. Sani, Z. Hanapi, & M. Abdullah, "An efficient anomaly intrusion detection method with feature selection and evolutionary neural network", Ieee Access, vol. 8, p. 70651-70663, 2020. https://doi.org/10.1109/ACCESS.2020.2986217.

[24] E. Li, "Incremental learning of lstm-autoencoder anomaly detection in three-axis cnc machines", 2023. https://doi.org/10.21203/rs.3.rs-3388986/v1.

[25] Z. Khan, "Network intrusion detection utilizing autoencoder neural networks", cana, vol. 31, no. 3s, p. 336-354, 2024. https://doi.org/10.52783/cana.v31.777.

[26] L. Shi, Y. Ma, Y. Lu, & L. Chen, "The application of computer intelligence in the cyber-physical business system integration in network security", Computational Intelligence and Neuroscience, vol. 2022, p. 1-10, 2022. https://doi.org/10.1155/2022/5490779.

[27] J. Seok, A. Kareem, & J. Hur, "Lstm-autoencoder for vibration anomaly detection in vertical carousel storage and retrieval system (vcsrs)", Sensors, vol. 23, no. 2, p. 1009, 2023. https://doi.org/10.3390/s23021009.

[28] F. Esmaeili, E. Cassie, H. Nguyen, N. Plank, C. Unsworth, & A. Wang, "Anomaly detection for sensor signals utilizing deep learning autoencoder-based neural networks", Bioengineering, vol. 10, no. 4, p. 405, 2023. https://doi.org/10.3390/bioengineering10040405.

[29] E. Khan, "Network intrusion detection using autoencode neural network", International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 10, p. 1678-1688, 2023. https://doi.org/10.17762/ijritcc.v11i10.8739.

[30] H. Nguyen, K. Tran, S. Thomassey, & M. Hamad, "Forecasting and anomaly detection approaches using lstm and lstm autoencoder techniques with the applications in supply chain management", International Journal of Information Management, vol. 57, p. 102282, 2021. https://doi.org/10.1016/j.ijinfomgt.2020.102282 .

[31] M. Park, S. Chakraborty, Q. Vuong, D. Noh, J. Lee, J. Leeet al., "Anomaly detection based on time series data of hydraulic accumulator", Sensors, vol. 22, no. 23, p. 9428, 2022. https://doi.org/10.3390/s22239428.

[32] Y. Fu, Y. Du, Z. Cao, Q. Li, & X. Wang, "A deep learning model for network intrusion detection with imbalanced data", Electronics, vol. 11, no. 6, p. 898, 2022. https://doi.org/10.3390/electronics11060898.

[33] Z. Cheng, S. Wang, P. Zhang, S. Wang, X. Liu, & E. Zhu, "Improved autoencoder for unsupervised anomaly detection", International Journal of Intelligent Systems, vol. 36, no. 12, p. 7103-7125, 2021. https://doi.org/10.1002/int.22582.

[34] H. Arab, I. Ghaffari, R. Evina, S. Tatu, & S. Dufour, "A hybrid lstm-resnet deep neural network for noise reduction and classification of v-band receiver signals", Ieee Access, vol. 10, p. 14797-14806, 2022. https://doi.org/10.1109/ACCESS.2022.3147980.

[35] P. Park, P. Marco, H. Shin, & J. Bang, "Fault detection and diagnosis using combined autoencoder and long short-term memory network", Sensors, vol. 19, no. 21, p. 4612, 2019. https://doi.org/10.3390/s19214612.

[36] L. Wu and L. Ji, "Anomaly detection based on temporal convolution autoencoders", Journal of Physics Conference Series, vol. 2366, no. 1, p. 012041, 2022. https://doi.org/10.1088/1742-6596/2366/1/012041.

[37] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacheret al., "N-baiot—network-based detection of iot botnet attacks using deep autoencoders", Ieee Pervasive Computing, vol. 17, no. 3, p. 12-22, 2018. https://doi.org/10.1109/MPRV.2018.03367731.

[38] C. Tang, N. Luktarhan, & Y. Zhao, "An efficient intrusion detection method based on lightgbm and autoencoder", Symmetry, vol. 12, no. 9, p. 1458, 2020. https://doi.org/10.3390/sym12091458.

[39] A. Mirza and S. Coşan, "Computer network intrusion detection using sequential lstm neural networks autoencoders", p. 1-4, 2018. https://doi.org/10.1109/SIU.2018.8404689.

[40] M. Moukhafi, "Intelligent intrusion detection through deep autoencoder and stacked long short-term memory", International Journal of Electrical and Computer Engineering (Ijece), vol. 14, no. 3, p. 2908, 2024. https://doi.org/10.11591/ijece.v14i3.pp2908-2917.

[41] A. Singh and J. Jang-Jaccard, "Autoencoder-based unsupervised intrusion detection using multi-scale convolutional recurrent networks", 2022.

[42] O. Adeniyi, "Securing mobile edge computing using hybrid deep learning method", Computers, vol. 13, no. 1, p. 25, 2024. https://doi.org/10.3390/computers13010025.

[43] T. Vaiyapuri, "Deep self-taught learning framework for intrusion detection in cloud computing environment", Iaes International Journal of Artificial Intelligence (Ij-Ai), vol. 13, no. 1, p. 747, 2024. https://doi.org/10.11591/ijai.v13.i1.pp747-755.

[44] S. Wu, Q. Huang, & L. Zhao, "De-noising of transient electromagnetic data based on the long short-term memory-autoencoder", Geophysical Journal International, vol. 224, no. 1, p. 669-681, 2020. https://doi.org/10.1093/gji/ggaa424.

[45] M. Tawfik, "Optimized intrusion detection in iot and fog computing using ensemble learning and advanced feature selection", Plos One, vol. 19, no. 8, p. e0304082, 2024. https://doi.org/10.1371/journal.pone.0304082.

[46] J. Kang, C. Kim, J. Kang, & J. Gwak, "Anomaly detection of the brake operating unit on metro vehicles using a one-class lstm autoencoder", Applied Sciences, vol. 11, no. 19, p. 9290, 2021. https://doi.org/10.3390/app11199290.

[47] T. Kim, "An anomaly detection method based on multiple lstm-autoencoder models for in-vehicle network", Electronics, vol. 12, no. 17, p. 3543, 2023. https://doi.org/10.3390/electronics12173543.

[48] Y. Chen and C. Zhao, "Application of deep learning model in computer data mining intrusion detection", Applied Mathematics and Nonlinear Sciences, vol. 8, no. 2, p. 2131-2140, 2023. https://doi.org/10.2478/amns.2023.1.00318.

[49] A. Halbouni, T. Gunawan, M. Habaebi, M. Halbouni, M. Kartiwi, & R. Ahmad, "Machine learning and deep learning approaches for cybersecurity: a review", Ieee Access, vol. 10, p. 19572-19585, 2022. https://doi.org/10.1109/ACCESS.2022.3151248.

[50] E. Ramadanti, "Diabetes disease detection classification using light gradient boosting (lightgbm) with hyperparameter tuning", Sinkron, vol. 8, no. 2, p. 956-963, 2024. https://doi.org/10.33395/sinkron.v8i2.13530.

[51] M. Onoja, A. Jegede, N. Blamah, A. Olawale, & T. Omotehinwa, "Eemds: efficient and effective malware detection system with hybrid model based on xceptioncnn and lightgbm algorithm", Journal of Computing and Social Informatics, vol. 1, no. 2, p. 42-57, 2022. https://doi.org/10.33736/jcsi.4739.2022.

[52] M. Ferrag, Λ. Μαγλαράς, S. Moschoyiannis, & H. Janicke, "Deep learning for cyber security intrusion detection: approaches, datasets, and comparative study", Journal of Information Security and Applications, vol. 50, p. 102419, 2020. https://doi.org/10.1016/j.jisa.2019.102419.

[53] D. T and M. Saravanan, "An intellectual detection system for intrusions based on collaborative machine learning", International Journal of Advanced Computer Science and Applications, vol. 11, no. 2, 2020. https://doi.org/10.14569/IJACSA.2020.0110257.

[54] Eeday, S., S. Goteti, and S. P. Anne. "Experimental Investigation of Thermal Properties of Borassus Flabellifer Reinforced Composites and Effect of Addition of Fly Ash." International Journal of Engineering Trends and Technology 15.8 (2014): 379-382

[55] Susarla, Manoj, S. Harshavardhan, and Saranya Eeday. "Structural Analysis of the Drag Strut and Dynamic analysis of Aircraft Landing Gear." 1-8.

[56] Kosuri, S., & Eeday, S. (2024). Advancing Cybersecurity with Deep Learning: Innovative Approaches and Real-Time Applications. International Journal of Advanced Trends in Engineering and Management (IJATEM), 3(8), 52–62. ISSN (Online): 2583-7052. https://doi.org/10.59544/TBFO5911/IJATEMV03I08P5.

[57] S. Kosuri and S. Eeday, "Deep Learning for Network Intrusion Detection: A Study Using Convolutional Neural Networks," International Journal of Research in Electronics and Computer Engineering, vol. 12, no. 4, p. 10, 2024.

[58] Eeday and S. Kosuri, "Adaptive Network Intrusion Detection Using Asymmetric Deep Autoencoders with Benchmarking on Latest Cybersecurity Datasets," International Journal of Research in Electronics and Computer Engineering, vol. 12, no. 2, pp. 42–45, 2024.

[59] S. Eeday and S. Kosuri, "A Hybrid Federated Tree-Based Intrusion Detection System for Securing Autonomous Vehicles and V2X Networks Using Adaptive Boosting and Blockchain," International Journal of Research in Electronics and Computer Engineering, vol. 11, no. 4, pp. 38–41, 2023.

[60] Moraboena, S., Ketepalli, G., Ragam, P. (2020). A deep learning approach to network intrusion detection using deep autoencoder. Revue d'Intelligence Artificielle, Vol. 34, No. 4, pp. 457-463. https://doi.org/10.18280/ria.340410.

[61] M. Srikanth Yadav. and R. Kalpana., "Data Preprocessing for Intrusion Detection System Using Encoding and Normalization Approaches," 2019 11th International Conference on Advanced Computing (ICoAC), Chennai, India, 2019, pp. 265-269, https://doi.org/10.1109/ICoAC48765.2019.246851.

[62] M., Srikanth Yadav, and Kalpana R. "A Survey on Network Intrusion Detection Using Deep Generative Networks for Cyber-Physical Systems." Artificial Intelligence Paradigms for Smart Cyber-Physical Systems, edited by Ashish Kumar Luhach and Atilla Elçi, IGI Global, 2021, pp. 137-159. https://doi.org/10.4018/978-1-7998-5101-1.ch007.

[63] Srikanth yadav M., R. Kalpana, Recurrent nonsymmetric deep auto encoder approach for network intrusion detection system, Measurement: Sensors, Volume 24, 2022, 100527, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2022.100527.

[64] Srikanth Yadav, M., Kalpana, R. (2022). Effective Dimensionality Reduction Techniques for Network Intrusion Detection System Based on Deep Learning. In: Jacob, I.J., Kolandapalayam Shanmugam, S., Bestak, R. (eds) Data Intelligence and Cognitive Informatics. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-6460-1_39.

[65] Gayatri, K., Premamayudu, B., Yadav, M.S. (2021). A Two-Level Hybrid Intrusion Detection Learning Method. In: Bhattacharyya, D., Thirupathi Rao, N. (eds) Machine Intelligence and Soft Computing. Advances in Intelligent Systems and Computing, vol 1280. Springer, Singapore. https://doi.org/10.1007/978-981-15-9516-5_21.

[66] Yadav, M. Srikanth, K. Sushma, and K. Gayatri. "Enhanced Network Intrusion Detection Using LSTM RNN." International Journal of Advanced Science and Technology 29.5 (2020): 7210-7220.

[67] Patil, A., and S. Yada. "Performance analysis of anomaly detection of KDD cup dataset in R environment." Int. J. Appl. Eng. Res. 13.6 (2018): 4576-4582.

[68] Saheb, M.C.P., Yadav, M.S., Babu, S., Pujari, J.J., Maddala, J.B. (2023). A Review of DDoS Evaluation Dataset: CICDDoS2019 Dataset. In: Szymanski, J.R., Chanda, C.K., Mondal, P.K., Khan, K.A. (eds) Energy Systems, Drives and Automations. ESDA 2021. Lecture Notes in Electrical Engineering, vol 1057. Springer, Singapore. https://doi.org/10.1007/978-981-99-3691-5_34.

[69] Patil, A., and M. Srikanth Yadav. "Performance analysis of misuse attack data using data mining classifiers." International Journal of Engineering & Technology 7.4 (2018): 261-263. https://doi.org/10.14419/ijet.v7i4.36.23782.

[70] R. Padmaja and P. R. Challagundla, "Exploring A Two-Phase Deep Learning Framework for Network Intrusion Detection," 2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2024, pp. 1-5, https://doi.org/10.1109/SCEECS61402.2024.10482198.

[71] Bhuyan, H.K., Ravi, V. & Yadav, M.S. multi-objective optimization-based privacy in data mining. Cluster Comput 25, 4275–4287 (2022). https://doi.org/10.1007/s10586-022-03667-3.