



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





An Enhanced AI-Based Network Intrusion Detection System using Generative Adversarial Networks

¹Shanmugapriya B ME, ²Thilagavathi A, ³Priyadharshini R, ⁴Swetha K, ⁵Dharani M

Assistant Professor, Department of CSE, Kavery Engineering, College (Autonomous), M.Kalipatti, Tamil Nadu, India

UG Students, Department of CSE, Kavery Engineering College (Autonomous), M. Kalipatti, Tamil Nadu, India

ABSTRACT: A Network Intrusion Detection System (NIDS) is an essential security element that continuously inspects network traffic for harmful or undesirable packets in real-time. NIDS employs perspective analysis technology to scrutinize traffic patterns, recognize deviations from set baselines, and detect anomalies that could suggest malicious activity. Training effective NIDS models is challenging due to the difficulty of acquiring sufficient data. The rapid occurrence of massive intrusion incidents on computer networks has made network security vulnerable. To solve this issue, a Generative Adversarial Network (GAN) model is suggested to enhance the performance of NIDS. Furthermore, the preprocessing used the Min- Max Normal Scaling (MMNS) technique to evaluate the data analysis, feature detection, and modeling. Next, the Correlation-Based Bat Algorithm (CBA) is introduced to select optimal features for dimensionality reduction based on correlations between features. Finally, integrating GAN technology into NIDS will improve intrusion detection performance for attacks with limited training data. The suggested GAN algorithm is capable of identifying and enhancing NIDS attacks by utilizing performance metrics like recall, precision, time complexity, and accuracy.

KEYWORDS: Artificial intelligence, NIDS, CIC-IDS-2017 dataset, GAN, CBA, and MMNS and feature selection.

I. INTRODUCTION

With the rapid expansion of digital infrastructure, cybersecurity threats are outpacing traditional network security. Network Intrusion Detection Systems (NIDS) are crucial for detecting malicious activity, but signature and rule-based methods struggle with evolving attacks, scalability, and high false positive rates. Consequently, Artificial Intelligence (AI) is being leveraged to improve NIDS through adaptive and intelligent threat detection. AI empowers NIDS to learn from data, detect subtle anomalies, and generalize to identify novel attacks. LENS-XAI, for instance, is a lightweight annotation framework that uses difference autoencoders and knowledge processing to balance model performance and analysis in low- resource environments [1].

By focusing on key features, XAI has also been used for feature selection to increase the expressiveness of models and enhance classifier performance [2]. Further developments in self-supervised learning systems have addressed the absence of labeled data. By converting the network data into an image format, the computer table enables a robust blind autoencoder to identify irregularities without requiring supervised training [3]. Numerous security mechanisms have been introduced to address these adversarial threats. However, many of them fail to achieve a balanced approach among robust resistance to attacks, high detection accuracy on untainted data, and preserving the integrity of traffic flow operations [4-5].

Attackers might covertly change hostile traffic signatures to avoid detection and perhaps breach critical systems if they are not aware of the precise mechanisms of the target NIDS. Protecting against these hostile attacks is difficult, but not impossible.

This paper mainly improves the efficiency of NIDS by proposing a GAN model. Furthermore, the accuracy can be enhanced using the CIC-IDS-2017 dataset from Kaggle for training and testing. The preprocessing step uses the MMNS technique for data analysis, feature detection, and modeling. A CBA method is introduced to select the optimal features for dimensionality reduction based on the correlations between the features.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SURVEY

A joint intrusion detection technique that uses unsupervised learning minimizes the need for labeling. By using a novel [6] combinatorial K-means clustering initialization, they ensure privacy without offering efficiency. Explore various techniques for NIDS, as many have noted that there is less emphasis on signature-based or hybrid approaches and more on anomaly-based IDS [7]. Offers a self-supervised online Deep Learning (DL) framework, removing the necessity for labeled data. The system employs an automated deep probabilistic neural network capable of adapting to network traffic in real-time [8]. The Adaptive Neuro- Fuzzy Inference System (ANFIS) intrusion detection ensemble analysis integrates neural networks and fuzzy logic [9]. Their system showed enhanced adaptability and accuracy in classifying network traffic. DL models are combined with a Support Vector Machine (SVM) algorithm to enhance IDS capabilities [10]. Furthermore, the hybrid approach increases detection accuracy across various datasets to achieve significant improvements.

The IDS detection is based on a deep Q-network in a reinforcement learning framework that dynamically adapts to new threats and improves the system's resilience [11]. The author[12] has introduced ensemble voting classifiers that combine XGBoost, random forests, and decision trees. This DRX-based approach achieves greater accuracy and precision on datasets such as NSL-KDD and CIC-IDS2017. A study [13] used a combination of a random forest classifier and permutational feature importance to identify salient network traffic features. This method achieves a weighted F1 score of 89.8% on the CICIDS-2017 dataset. Optimal feature selection with the hybrid DL model demonstrated optimal accuracy on several datasets, including UNSW-NB15 and NSL-KDD. This architecture reduces the computing cost while efficiently handling the class imbalance issue [14]. Assessed supervised learning algorithms, focusing on K-Nearest Neighbors (KNN) to detect intrusion attempt anomalies [15]. Their results underscore the effectiveness of these models in enhancing network security.

The [16] novel machine learning (ML) model combines feature extraction, layered feature embedding, and oversampling approaches to tackle the problems of massive and imbalanced datasets. The approach's accuracy surpasses 79.9% on datasets like CIC-IDS2017.

A study [17] presented a hybrid feature selection approach integrating algorithms like CatBoost and XGBoost with stacked ensemble learning. The model enhances multi-class classification performance across benchmark datasets. The detailed analysis compares seven machine learning algorithms on network flow datasets. This study [18] provides insight into each approach's strengths and challenges. This paper [19] introduces a DL method for forecasting network intrusion alerts.

It also presents a Gated Recurrent Units (GRU) capable of learning dependencies within security alert sequences. The integrated security architecture may enhance ML-based NIDS resilience against adversarial prevention [20]. Validation with NSL-KDD and UNSW-NB15 datasets shows a significant improvement in ML-based NIDS, achieving a 35% increase in detection accuracy and a 12.5% decrease in false positives.

III. PROPOSED METHODOLOGY

This section examined and identified attacks and benign traffic utilizing the proposed GAN model, based on the efficiency of NIDS. Additionally, improvements were attained through data preprocessing, feature selection, and classification based on features from the CICIDS-2017 dataset sourced from Kaggle.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

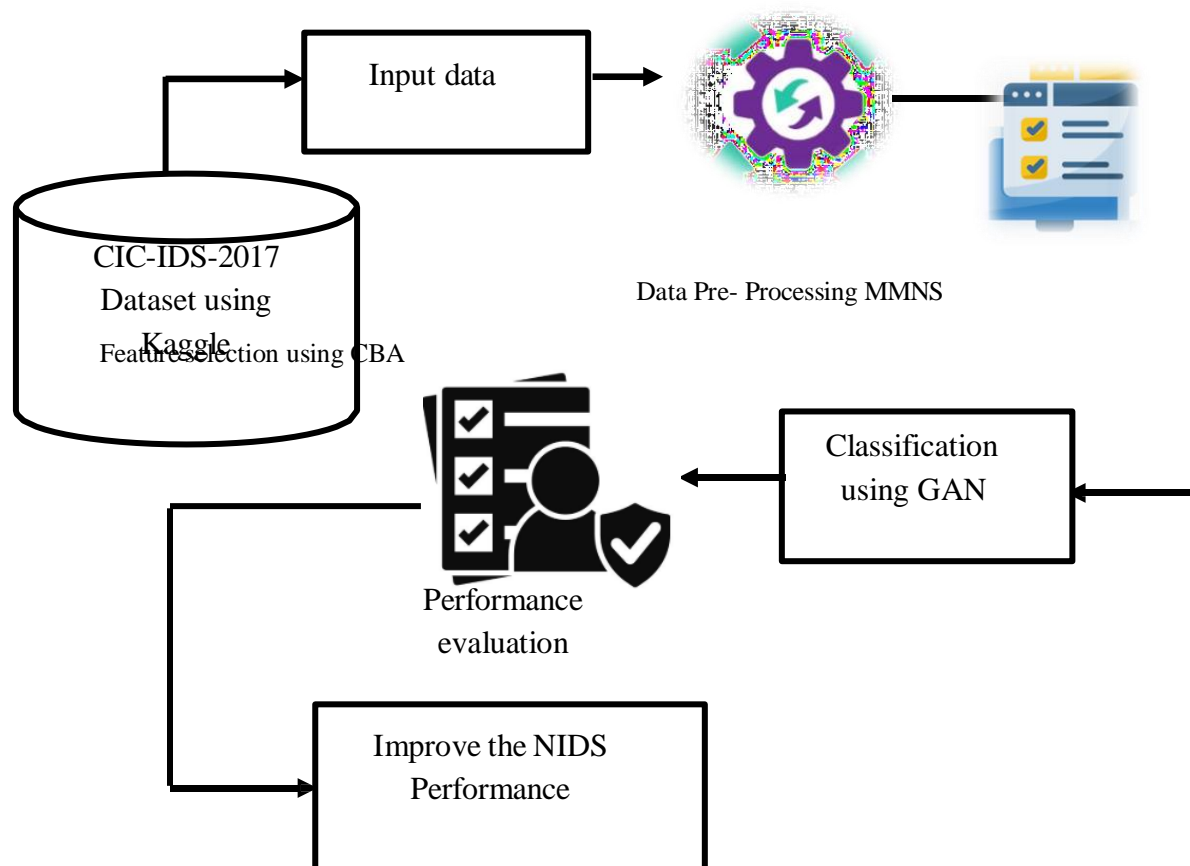


Figure 1. The Proposed GAN Method Architecture Diagram

As illustrated in Figure 1, the performance and accuracy of NIDS are improved by implementing the GAN model. To accurately evaluate and refine the proposed system, the widely recognized CIC-IDS-2017 dataset from Kaggle is used for training and testing to provide a standardized and relevant environment for performance evaluation. A key component of the approach is a comprehensive pre-processing stage using MMNS techniques, which enables robust data analysis, accurate feature detection, and effective model generation. Furthermore, the CBA method is strategically introduced to address the challenges caused by high dimensionality and redundant features. The proposed CBA approach can intelligently identify and select the best feature subset that effectively reduces dimensionality while retaining the most relevant information for accurate NIDS.

3.1 CIC-IDS2017 Dataset Collection

This section utilizes the CIC-IDS2017 dataset from Kaggle for evaluating and benchmarking NIDS and network security solutions. Additionally, the dataset is applied to assess various types of network activity, encompassing both attacks and benign traffic. A comprehensive analysis of the widespread intrusion is conducted using the CIC-IDS2017 dataset at <https://www.kaggle.com/datasets/chethuhn/network-intrusion-dataset>.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Destination Port	Flow Duration	Total Fwd Packets	Total Backward Packets	Total Length of Fwd Packets	Total Length of Bwd Packets	Packet Length	Fwd Packet Length Min	Fwd Packet Length Mean
54865	3	2	0	12	0	6	6	6
55054	109	1	1	6	6	6	6	6
55055	52	1	1	6	6	6	6	6
46236	34	1	1	6	6	6	6	6
54863	3	2	0	12	0	6	6	6
54871	1022	2	0	12	0	6	6	6
54925	4	2	0	12	0	6	6	6
54925	42	1	1	6	6	6	6	6
9282	4	2	0	12	0	6	6	6
55153	4	2	0	37	0	31	6	18.5
55143	3	2	0	37	0	31	6	18.5
55144	1	2	0	37	0	31	6	18.5
55145	4	2	0	37	0	31	6	18.5
55254	3	3	0	43	0	31	6	14.33333333
36206	54	1	1	0	0	0	0	0
53524	1	2	0	0	0	0	0	0
53524	154	1	1	0	0	0	0	0
53526	1	2	0	0	0	0	0	0

Figure 2. Dataset Feature Collection

As shown in Figure 2, using the CIC-IDS2017 dataset, a comprehensive analysis of a wide-ranging attack was conducted to collect malicious and normal network activity features. 1,197,856 instances of the dataset's 225,746 unique features were used for training, while 27,890 examples were reserved aside for testing.

3.2 Min-Max Normalized Scaling (MMNS)

This section describes the preprocessing applied in the MMNS technique for evaluating data analysis, feature detection, and modeling. The accuracy of anomaly detection hinges on its capacity to manage non-Gaussian feature distributions instead of relying on normalized scaling methods, which is achieved through the implementation of the MMNS method. Additionally, the signature-based strategy of NIDS enables us to regularize data from the Min- Max technique by refining the path loss function using features within the dataset.

Output values are computed as the minimum and maximum values of the column between 0 and 1, as indicated in Equation 1. Let's assume x —original dataset value, x_{max} —largest number of columns, x_{min} —shortest number of columns.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

The MMNS model and learning rate optimization can offer a global or local min-max normality estimate to provide quick and precise NIDS detection.

3.3 Correlation-Based Bat Algorithm (CBA)

The CBA provides a one-dimensional reduction method that strategically identifies and retains the least correlated features. This CBA method enhanced performance and interpretability by minimizing redundancy and addressing high-dimensional NIDS CIC-IDS- 2017 datasets. A CBA-based feature selection method evaluates the importance and relevance of the selected feature subsets in the CIC-IDS2017 dataset. Additionally, a CBA approach can be used to select the best features and assess the fitness function and completeness of the reduced feature subset.

Algorithm

Input: CIC-IDS2017 Normalized data x'

Output: Optimal subset feature selection X_{Best}



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Start

1. Estimate the number of bats
- $X_i = (X_{i1}, \dots, X_{id})(i = 1, 2, \dots, n)$ and v_i
2. Calculate frequency, pulse rate and noise
3. If $1 \leq t \leq \text{max no of iteration}$, do
4. For each $i = 1$ to n , do
5. Create a new frequency f_i
6. Update the Rules for Position and Speed x_i and v_i
7. If pulse rate $r^t < \text{Rand}(0,1)$ then
8. Select a x_i from X_{best}
9. End if
10. End for each
11. End if
12. Return select subset features X_{Best}
13. End

The CBA can select optimal features and evaluate the fitness and completeness of the reduced feature subset.

3.4 Generative Adversarial Networks (GANs)

This section will integrate GAN technology and use the CICIDS-2017 dataset for training and testing to improve the performance of NIDS attack detection. It uses the GAN algorithm to analyze the generator and discriminator, which consist of two neural networks. A generator network enables the creation of data features from a given dataset. During the training process, the GAN algorithm improves the discrimination ability of the generator and discriminator networks. Furthermore, GAN methods can detect attacks to determine malicious and normal network activities.

The output produced by a GAN represents the probability that the input is real. Compute the parser's loss function, as shown in equation 2. Let's assume L_D – loss discriminator's,

(x) – produces an output, $-E_{x \sim p_{\text{data}}}(x) [\log D(x)]$ – expected value of the logarithm,

$E_{x \sim p_{\text{ge}}}(x) [\log(1 - D(x))]$ – discriminator's output for generated data.

$$L_D = -E_{x \sim p_{\text{data}}}(x) [\log D(x)] - E_{x \sim p_{\text{gen}}}(x) [\log(1 - D(x))] \quad (2)$$

The generator's objective is to minimize the probability of correctly classifying the data generated by the parser as spurious. Calculate the generator's loss function, as demonstrated in equation 3. Let's assume L_G – Generating loss, x – Data samples, p_{data} – Real data distribution, p_{gen} – Generated data distribution.

$$L_G = -E_{x \sim p_{\text{ge}}}(x) [\log(1 - D(x))] \quad (3)$$

The generator is assessed to reduce the anticipated logarithm of the output based on the generated data.

IV. RESULT AND DISCUSSION

This section implemented the features on the test and training datasets to optimize the NIDS performance metrics. Furthermore, the NIDS performance metrics, such as precision, recall, time complexity, precision, and F-measure, can be calculated using the confusion matrix parameters. Compared with traditional techniques such as SVM, KNN, and XGBoost, the proposed GAN technique for verifying the accuracy of NIDS achieves the highest improvement in test and training accuracy performance metrics.

10.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Table 1 Simulation Parameter

Simulation	Value
Name dataset	Network Intrusion dataset (CIC-IDS- 2017)
No of Dataset	225746
Training	197856
Testing	27,890
Language	Python
Tool	Jupyter

As shown in Table 1, the simulation parameters have been implemented and evaluated using the Python-based Jupyter tool based on CIC-IDS-2017.

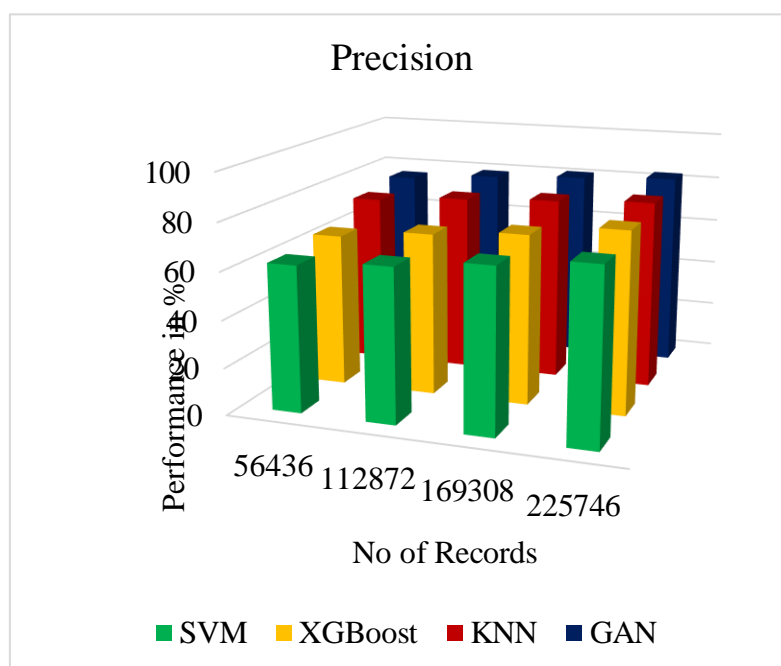


Figure 3. Analysis of Precision

The precision performance measure can enhance the effectiveness of the proposed method through the NIDS function. Figure 3 illustrates that the precision analysis indicates the GAN method achieves an 83% rate, outperforming the earlier SVM, KNN, and XGBoost methods. Additionally, the proposed technique scores 73%, 77%, and 80% when compared to the previous approaches, marking a notable enhancement in the precision analysis of the NIDS.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

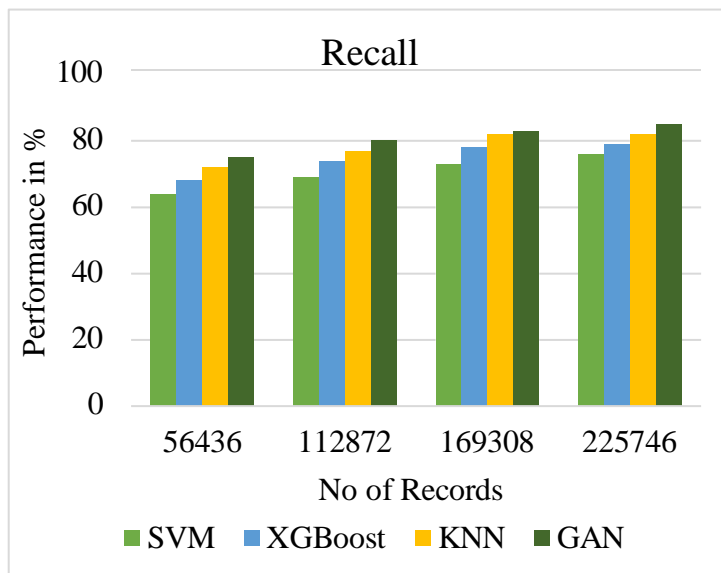


Figure 4. Analysis of Recall

The recall performance measurement can increase the effectiveness of the proposed method through NIDS capabilities. Figure 4 shows that the recall analysis indicates that the GAN method outperforms the previous SVM, KNN, and XGBoost methods, reaching a rate of 85%. Also, the proposed method shows a significant improvement in the recall analysis of NIDS, achieving scores of 76%, 79%, and 82% compared to previous approaches.

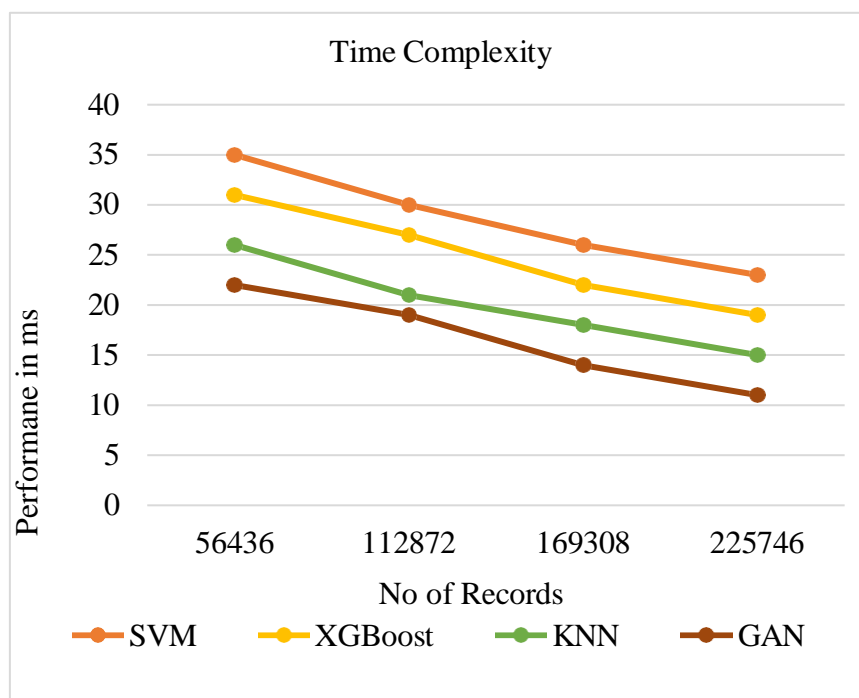


Figure 5. Analysis of Time Complexity

The time complexity performance measurement can increase the effectiveness of the proposed method through NIDS capabilities. Figure 5 illustrates that the time complexity



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

analysis indicates that the GAN method outperforms the previous SVM, KNN, and XGBoost methods, reaching a rate of 11.02ms. Furthermore, the proposed method shows a significant improvement in the time complexity analysis of NIDS, achieving scores of 26ms, 20ms, and 15ms compared to previous approaches.

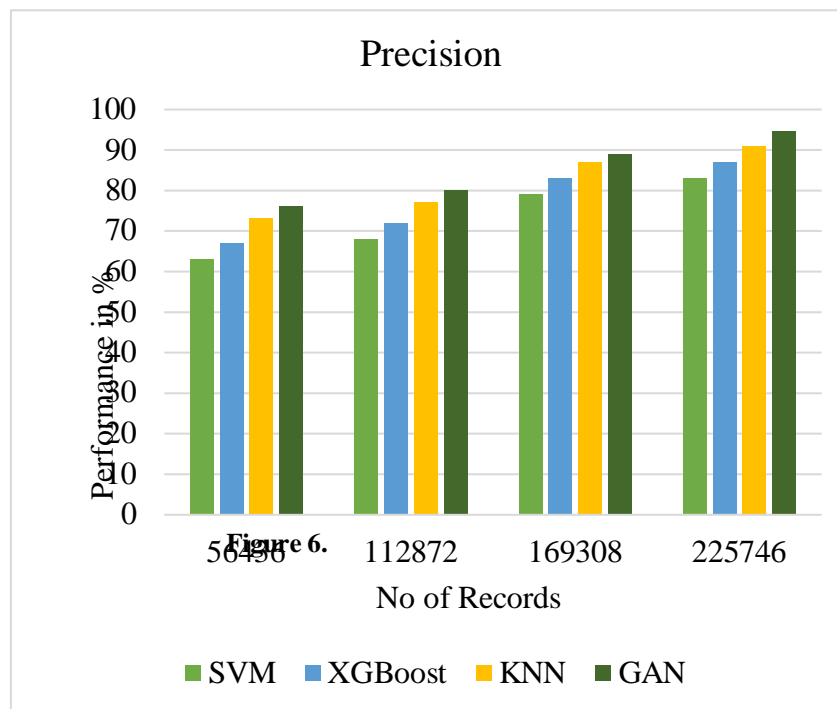


FIGURE6: Analysis of Accuracy

Accuracy performance measurement can enhance the effectiveness of the suggested method using NIDS capabilities. As indicated in Figure 5, accuracy analysis reveals that the GAN method surpasses the earlier SVM, KNN, and XGBoost methods, achieving a rate of 94.6%. Additionally, the proposed method demonstrates a notable improvement in NIDS accuracy analysis, attaining scores of 83%, 87%, and 91% when compared to prior approaches.

V. CONCLUSION

In conclusion, implementing a GAN model enhances NIDS performance and accuracy. The CIC-IDS-2017 dataset from Kaggle is used for training and testing to provide a standardized evaluation environment. A comprehensive pre-processing stage using MMNS techniques enables robust data analysis, accurate feature detection, and effective model generation. Furthermore, a CBA method addresses challenges from high dimensionality and redundant features by intelligently selecting the best feature subset, reducing dimensionality while preserving relevant information for accurate NIDS. The GAN method showcases superior accuracy, as evidenced by a detailed analysis of its performance. Specifically, this analysis highlights that the GAN method outperforms several established ML techniques, namely SVM, KNN, and XGBoost. The performance margin is significant, with the GAN model reaching an impressive accuracy rate of 94.6%. Furthermore, the proposed methodology demonstrates a substantial and noticeable improvement in Network NIDS accuracy analysis. These comparisons demonstrate that the GAN method achieves accuracy scores of 83%, 87%, and 91%, indicating a positive improvement in NIDS performance compared to prior methodologies. This increased accuracy highlights the potential of GANs for enhancing network security and detection.



International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

REFERENCES

1. Yagiz, A. I., & Goktas, F. (2025). LENS-XAI: Redefining Lightweight and Explainable Network Security through Knowledge Distillation and Variational Autoencoders for Scalable Intrusion Detection in Cybersecurity, [arXiv:2501.00790](https://arxiv.org/abs/2501.00790)
2. Arreche, O., Guntur, T., & Abdallah, M. (2024). XAI-based Feature Selection for Improved Network Intrusion Detection Systems. [ArXiv. https://arxiv.org/abs/2410.10050](https://arxiv.org/abs/2410.10050).
3. Li, E., Shang, Z., Gungor, O., & Rosing, T. (2025). SAFE: Self-Supervised Anomaly Detection Framework for Intrusion Detection. [ArXiv. https://arxiv.org/abs/2502.07119](https://arxiv.org/abs/2502.07119).
4. Awad, Z., Zakaria, M., & Hassan, R. (2025). An enhanced ensemble defense framework for boosting adversarial robustness of intrusion detection systems. *Scientific Reports*, 15(1), 1-23. <https://doi.org/10.1038/s41598-025-94023-z>
5. Zhang, C., Costa-Perez, X. & Patras, P. Adversarial attacks against deep learning-based network intrusion detection systems and defense mechanisms. *IEEE/ACM Trans. Networking*. **30** (3), 1294–1131. <https://doi.org/10.1109/TNET.2021.3137084> (2022).
6. Gourceyraud, M., Salem, R. B., Neal, C., Cuppens, F., & Cuppens, N. B. (2025). Federated Intrusion Detection System Based on Unsupervised Machine Learning. [ArXiv. https://arxiv.org/abs/2503.22065](https://arxiv.org/abs/2503.22065)
7. Abdulganiyu OH, Tchakoucht TA, Saheed YK (2024) Towards an efficient model for network intrusion detection system (ids): systematic literature review. *Wirel Netw* 30(1):453–482.
8. Nakip, M., & Gelenbe, E. (2023). Online Self-Supervised Deep Learning for Intrusion Detection Systems. [ArXiv. https://doi.org/10.1109/TIFS.2024.3402148](https://doi.org/10.1109/TIFS.2024.3402148)
9. J. Sharma, Sonia, K. Kumar, P. Jain, R. H. C. Alfilh, & H. Alkattan, Trans, "Enhancing Intrusion Detection Systems with Adaptive Neuro-Fuzzy Inference Systems," (2025). *MesopotamianJournalCyberSecurity*,5(1), 1-10. <https://doi.org/10.58496/MJCS/2025/001>.
10. Latif, N., Ma, W. & Ahmad, H.B. Advancements in securing federated learning with IDS: a comprehensive review of neural networks and feature engineering techniques for malicious client detection. *Artif Intell Rev* **58**, 91 (2025). <https://doi.org/10.1007/s10462-024-11082-w>
11. Chakrawarti , A, & Shrivastava, S. S. (2024). "Enhancing Intrusion Detection System using Deep Q-Network Approaches based on Reinforcement Learning," *International Journal of Intelligent Systems and Applications in Engineering*, 12(12s), 34–45. Retrieved from <https://www.ijisae.org/index.php/IJISAE/article/view/4490>.
12. Farooqi, A. H., Akhtar, S., Rahman, H., Sadiq, T., & Abbass, W. (2023). Enhancing Network Intrusion Detection Using an Ensemble Voting Classifier for Internet of Things. *Sensors*, 24(1), 127. <https://doi.org/10.3390/s24010127>.
13. Abdelaziz, M.T., Radwan, A., Mamdouh, H. *et al.* Enhancing Network Threat Detection with Random Forest-Based NIDS and Permutation Feature Importance. *J Netw Syst Manage* **33**, 2 (2025). <https://doi.org/10.1007/s10922-024-09874-0>
14. Kumar, N., & Sharma, S. (2022). A Hybrid Modified Deep Learning Architecture for Intrusion Detection System with Optimal Feature Selection. *Electronics*, 12(19), 4050. <https://doi.org/10.3390/electronics12194050>
15. Talukder, M. A., Islam, M. M., Uddin, M. A., Hasan, K. F., Sharmin, S., Alyami, S. A., & Moni, M. A. (2024). Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction. *Journal of Big Data*, 11(1), 1-44. <https://doi.org/10.1186/s40537-024-00886-w>.
16. Huang, Z.; Li, Z.; Zhang, J. (2023). Enhancing network security through machine learning: A study on intrusion detection system using supervised algorithms. *Applied and Computational Engineering*,19,50-66.
17. Alsaffar, A. M., & Zolbanin, H. M. (2024). Shielding networks: Enhancing intrusion detection with hybrid feature selection and stack ensemble learning. *Journal of Big Data*, 11(1), 1-32. <https://doi.org/10.1186/s40537-024-00994-7>.
18. Mondragon, J.C., Branco, P., Jourdan, GV. *et al.* Advanced IDS: a comparative study of datasets and machine learning algorithms for network flow-based intrusion detection systems. *Appl Intell* **55**, 608 (2025). <https://doi.org/10.1007/s10489-025-06422-4>.
19. Ansari MS, Bartoš V, Lee B (2022) GRU-based deep learning approach for network intrusion alert prediction. *Future Gener Comput Syst* 128:235–247. <https://doi.org/10.1016/j.future.2021.09.040>
20. Tafreshian, B., & Zhang, S. (2025). A Defensive Framework Against Adversarial Attacks on Machine Learning-Based Network Intrusion Detection Systems. [ArXiv. https://arxiv.org/abs/2502.15561](https://arxiv.org/abs/2502.15561).



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  ijircce@gmail.com



www.ijircce.com

Scan to save the contact details