

SDAE-GAN for Handling Class Imbalance in the UNSW-NB15 Dataset

Sebastianus Hartantyo, Alva Hendi Muhammad

Master of Information Technology, Universitas Amikom Yogyakarta, INDONESIA

E-mail: hartantyo@students.amikom.ac.id, alva@amikom.ac.id

| Submitted: February 25, 2026 | Revised: March 05, 2026 | Accepted: May 30, 2026 |

| Published: October 23, 2026 |

ABSTRACT

Dataset for Intrusion Detection Systems (IDS) are often 'imbalanced', where normal data conditions are too abundant while attack data is very scarce. This condition makes machine learning performance in detecting attacks poor, especially for attacks that are indeed rare. Previous research addressed this issue by applying Stacked Denoising Autoencoder (SDAE), which can reduce the dimensions of features in large datasets. In this study, researchers explored the potential use of SDAE but combined it with Generative Adversarial Networks (GAN). SDAE was adopted not for dimension reduction, but with a different approach to augmentation or multiplying minority attack class data. With the SDAE-GAN hybrid model, the researchers proposed an architecture using an SDAE-based generator trained in a GAN framework to generate realistic synthetic data for each minority attack class in the UNSW-NB15 dataset. This process balanced the data distribution before training the classification model. The researchers then tested the model using Support Vector Machine (SVM) on the balanced data. The results were very significant, with the SVM model achieving 97% accuracy, an F1-Score of 98.44%, and, most importantly, a near-perfect recall value of 99.67% in detecting attack classes. This shows that the SDAE-GAN hybrid augmentation method can overcome the imbalance problem and drastically improve the IDS's ability to detect attacks from minority classes.

Keywords: intrusion detection system, class imbalance, data augmentation, stacked denoising autoencoder, generative adversarial network, UNSW-NB15, SVM.

INTRODUCTION

In today's modern digital era, network security has become one of the most fundamental and indispensable aspects. The increasingly widespread digital infrastructure has led to more diverse and complex cyber threats. Therefore, the Intrusion Detection System (IDS) plays a key role as the main line of defense in detecting abnormal activities that could potentially harm the network. The use of machine learning (ML) and deep learning in IDS development is not new; various studies have shown that machine learning-based approaches can significantly improve detection accuracy[1]. However, there is one classic problem that remains a major challenge, namely class imbalance in network traffic datasets.

IDS datasets such as UNSW-NB15 generally have a very skewed distribution between normal data and attack data. Normal data is far more dominant, while attack samples, especially the latest or rare types of attacks, are very minimal. This imbalance causes serious problems in the model training process. Machine learning models trained on such datasets tend to be biased, focusing more on recognizing normal patterns and ignoring attack patterns [2]. As a result, even though the model appears to have high accuracy, in reality, it often fails to recognize minority attacks, which are often the most dangerous and most important attacks for IDS to detect.

Various efforts have been made to overcome the problem of dataset imbalance. One previous study successfully utilized the Stacked Denoising Autoencoder (SDAE) approach to reduce dimensions and extract features in the NSL-KDD dataset. The results of this study showed a significant improvement in classification performance [3]. In conclusion, the researchers recommended further development in the form of integrating SDAE with generative models such as Generative Adversarial Network (GAN) to overcome class imbalance more effectively.

This study emerged as a follow-up and development of this idea, but with a different approach. Instead of using SDAE only for dimension reduction, this study adapted the SDAE architecture to be used as a Generator within the GAN framework. In this way, the proposed SDAE-GAN hybrid model is specifically designed to generate high-quality synthetic data as a solution for augmenting imbalanced datasets. The main objective is to improve the representation of minority classes in the dataset, so that the classification model has a more balanced learning opportunity and is able to better understand rare attack patterns.

The hypothesis is that after augmentation using the SDAE-GAN hybrid, the performance of the classification model will improve significantly, especially in recognizing attacks that were previously difficult to detect. To prove this hypothesis, experiments were conducted using the UNSW-NB15 dataset and the performance results were evaluated using the Support Vector Machine (SVM) classification method.

RESEARCH METHODS

The research methodology began with the data preparation stage as the main foundation of the experiment process. In the first step, the original dataset was divided into two main parts, namely the dataset for the training process and the dataset for testing, so as to ensure that the model built could be tested objectively using data that had never been seen before. After the data division stage was completed, the process continued with the application of the SDAE-GAN hybrid model, which functions as a data augmentation technique to handle class imbalance and enrich the sample distribution in the dataset. This model is designed to generate high-quality synthetic data that is representative of rarely found attack patterns. The final stage of this methodology is training and evaluating the classification model using the Support Vector Machine (SVM) algorithm. The evaluation is carried out with the aim of assessing the model's performance in detecting attack and normal classes more accurately after data augmentation. With this series of stages, this study aims to provide more comprehensive results in improving the performance of intrusion detection systems.

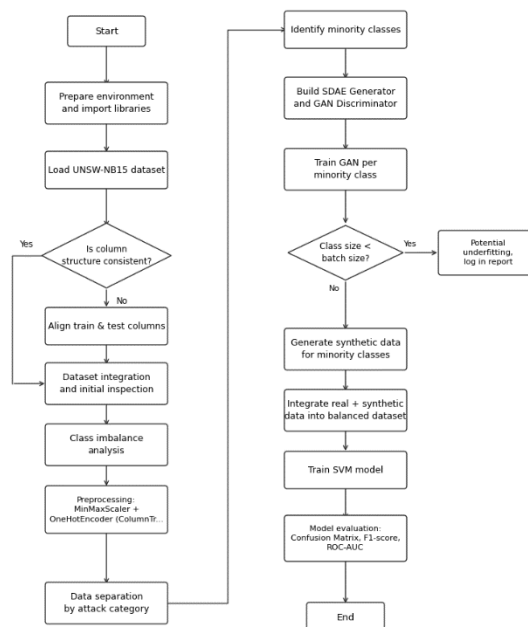


Figure 1. SDAE-GAN Hybrid Methodology

Dataset

The dataset used in this study is the UNSW-NB15 dataset, which is currently one of the modern benchmarks for IDS research. This dataset was created by the Australian Centre for Cyber Security (ACCS) and contains a mixture of normal traffic and nine types of the latest attacks [1]. The UNSW-NB15 dataset is considered capable of describing current network traffic conditions more realistically. It contains not only a variety of complex traffic characteristics, but also various types of attacks that are commonly encountered today. Overall, this dataset records nine main categories of attacks, such as Fuzzers, Analysis, Backdoor, DoS, Exploits, Generic, Reconnaissance, Shellcode, and Worms [2].

Table 1. Attack Distribution Of The UNSW-NB15 Dataset

No	Attack Type	Number of Samples
1.	Generic	5887
2.	Exploits	44,525
3.	Fuzzers	24246
4.	DoS	16353
5.	Reconnaissance	13987
6.	Analysis	2677
7.	Backdoor	2329
8.	Shellcode	1511
9.	Worms	174

An in-depth analysis of this data distribution reveals serious challenges in the methodological approach. From the available data, the Normal class contains approximately 93,000 samples, while the Worms class only has 174 samples. This huge imbalance, with a ratio of approximately 1 to 534, is far beyond the ability of standard classification algorithms to work fairly. Without special handling, machine learning models tend to form biased decision boundaries, so that samples from rare classes such as Worms or Shellcode can be considered noise, rather than important signals that should be studied. Most studies that utilize this dataset but do not properly address the data imbalance issue tend to produce low detection performance in the Analysis, Backdoor, and Worms classes. A study conducted by Zoghi and Serpen (2024) revealed that class overlap and data distribution imbalance are the main factors contributing to the increased False Negative Rate in modern IDS models [3], [4].

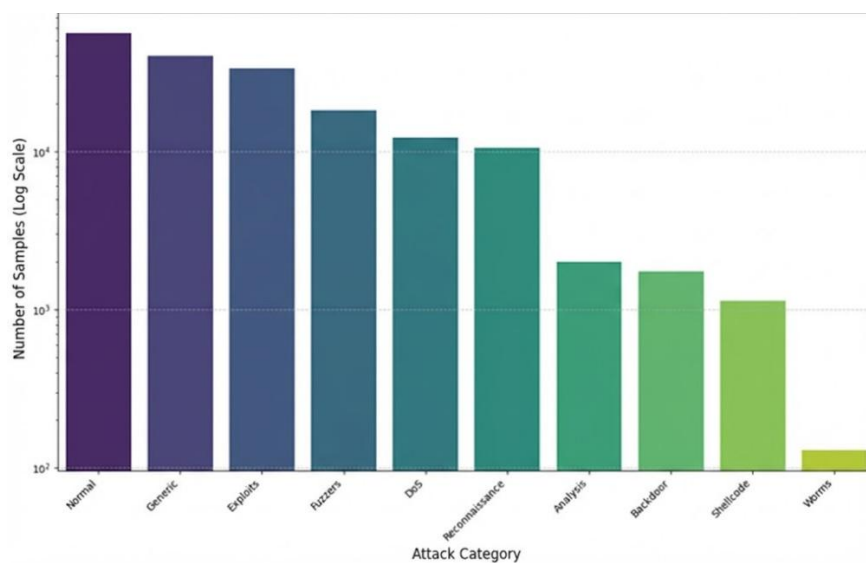


Figure 2. Class Distribution in The UNSW-NB15 Dataset

Data Preprocessing

Before the data is processed by the model, several important preparations must be made. The first preparation is to sort or identify features that are truly relevant, both categorical and numerical [5]. This process involves separating the UNSW-NB15 dataset into two large groups: majority classes such as *Normal*, *Generic*, and *Exploits*, and minority classes including *Analysis*, *Backdoor*, *DoS*, *Fuzzers*, *Reconnaissance*, *Shellcode*, and *Worms*. Each minority class is then trained using the SDAE-GAN module separately. This approach is crucial because training one large GAN model for all minority classes at once often fails due to mode competition [6]. By building a specific generator for each class, for example, a specific generator for *Worms* and another generator for *Shellcode*, the quality of the synthetic data produced can be optimally improved. After the synthetic data is successfully created, it is then combined with the original data to form a *Balanced Training Set*. Next, all numerical features are scaled so that their values are between 0 and 1. This is done to prevent features with large numbers from dominating the analysis results. Next, categorical features such as protocol, service, and state are converted into a numerical format that the model can understand. After the entire process is complete, the resulting dataset now has 189 features that are ready to be used and studied by the model.

The quality of the input data is a key factor in determining the success of a Deep Learning model. Therefore, the data preprocessing process is carried out carefully, starting with data cleaning to remove irrelevant features or those that only act as unique identifiers, such as IDs. Features with almost zero variance are also eliminated because they do not provide meaningful information. Missing values are treated using mean imputation for numerical data and mode imputation for categorical data to maintain the consistency of the data distribution structure. For categorical features, the One-Hot Encoding approach was chosen over Label Encoding, as the single numerical encoding method can cause hierarchical bias between categories. With One-Hot Encoding, each category is represented orthogonally, although this results in an increase in the number of features to 189. Categories that appear very rarely are combined into the "Others" label to control dimension explosion [7].

Meanwhile, numerical features are normalized using Min-Max Scaling to be within the range of 0.1. This normalization is important because the scale differences between features in the IDS dataset are extreme, for example, connection duration which is only a few seconds compared to sbytes values which can reach millions. Without normalization, features with large values can dominate loss calculations and update gradients, slowing convergence or trapping the model in suboptimal solutions. Min-Max Scaling was chosen because it is compatible with the activation functions used in SDAE and GAN architectures, such as Sigmoid and Tanh, which work optimally on limited value ranges. Thus, the learning process becomes more stable and efficient [8].

$$x^2 = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

Where:

- X : the original value of the feature before normalization
- min(x) : the minimum value of the feature
- max(x) : the maximum value of that feature
- x' : new value resulting from normalization that has been mapped to the range 0–1

SDAE-GAN Hybrid Model Architecture

This research adopts the GAN framework [9], which is essentially a 'game' between two models, namely the generator and the discriminator. The generator in the proposed model is based on Stacked Denoising Auto Encoder (SDAE). The generator is designed with an architecture inspired by deep autoencoder, where training begins by optimizing SDAE using data from the minority class. This step aims to capture stronger and more stable latent feature representations. Architecturally, this model consists of two main parts: the Encoder, which compresses 189-dimensional input data into a smaller latent space (e.g., from 128 to 64), and the Decoder, which then reconstructs the representation back to its original 189 dimensions. During the training stage, a *denoising* mechanism is applied, which distorts the original input x into $\tilde{x}$ by adding *masking noise* that randomly removes some values, for example with a probability of 20%. The SDAE model is then trained to minimize the *reconstruction error* between the output y and the original input x , so that the model learns to recover clean data even when given distorted input [10].

The *loss* function used is formulated as follows:

$$L_{SDAE} = \|x - y\|^2 + \lambda \Omega(W) \quad (2)$$

Where X

X : original data before noise is added

Y : reconstruction result from the model

$\|x - y\|^2$: the Mean Squared Error

Λ : regularization coefficient, determines how strong the penalty is

W : weights in the SDAE network

$\Omega(W)$: regularization function, often in the form of L2 regularization (sum of squared weights)

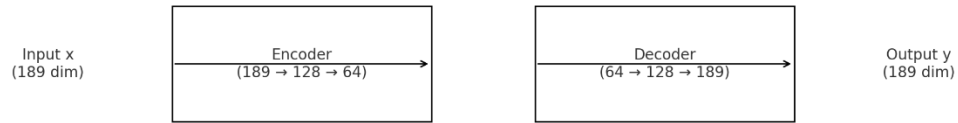


Figure 3. Visual Diagram of the SDAE Architecture

After the SDAE training process reaches convergence, the network weights in the Decoder section are extracted and used as the initialization for the Generator weights in the GAN model. At this stage, the Generator (G) receives input in the form of a random noise vector z derived from a standard normal distribution, then maps it to the data feature space. By utilizing the pre-trained SDAE weights, the Generator does not start learning from a completely random state, but already has the initial ability to represent attack characteristics. This strategy is an important component to prevent mode collapse [11].

Meanwhile, the Discriminator (D) acts as a conventional Multi-Layer Perceptron (MLP) network tasked with distinguishing between real data and synthetic data generated by G. Training is conducted adversarially, where the two networks compete with each other and are updated alternately. The optimized objective function is as follows:

Discriminator maximizes: $\log(D(x)) + \log(1 - D(G(z)))$

Generator minimizes: $\log(1 - D(G(z)))$

The training process continues until the Nash equilibrium point is reached, or until the quality of the synthetic data shows stability based on periodic evaluations. Because the UNSW-NB15 dataset has many minority classes, GAN training is not directly trained using the global dataset, but rather a separate GAN model is trained for each minority attack class. The process is carried out by first selecting one minority class, for example 'Worms'. Then, the proposed SDAE-GAN model is trained using the original 'Worms' data. The generator learns to mimic the 'Worms' data, while the discriminator learns to distinguish between real and fake 'Worms'. Once the model converges (is trained), the generator prints new synthetic 'Worms' samples. The model generates new data until the number of 'Worms' samples is equal to the number of samples in the majority class ('Normal'). This process is repeated for all other minority classes, resulting in a new balanced training dataset, a combination of original and synthetic data.

Classification and Evaluation

The dataset that has undergone augmentation and now has a balanced class distribution is then used to train the Support Vector Machine (SVM) model. At this stage, a Radial Basis Function (RBF) kernel is used, defined as:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (3)$$

To obtain the best performance, Hyperparameter Tuning is performed using Grid Search with 5-fold Cross Validation. The optimized parameters include the values of the C s (e.g., from 2^{-5} to 2^{15}) and the γ s (e.g., from 2^{-15} to 2^3). A larger C value is generally preferred to reduce classification errors in the training data, especially since the dataset is clean and balanced, while γ is used to control the extent to which each support

vector influences the decision boundary [12]. Model performance is evaluated on the test data using standard metrics such as Accuracy, Precision, Recall, F1-Score, and ROC AUC.

RESULT AND DISCUSSION

Augmentation Results

The augmentation process using the hybrid SDAE-GAN was successful. The class distribution in our training dataset, which was previously highly imbalanced, is now perfectly balanced. We successfully increased the number of samples in each minority attack category to be equivalent to the Normal class. In total, our balanced dataset has 560,000 samples, where each of the 10 classes (including 9 attack classes) has 56,000 samples.

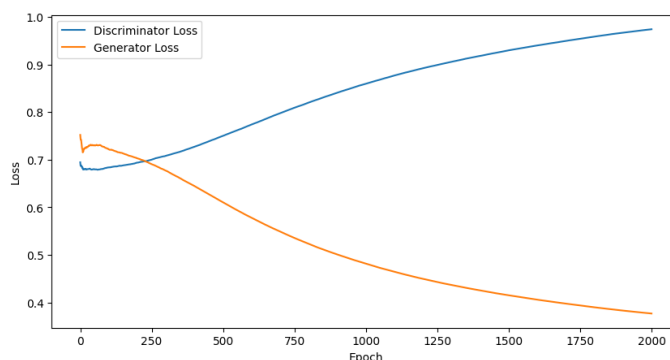


Figure 4. GAN Training Loss Curve For "*Worms*"

The quality of the synthetic data generated by the SDAE-GAN hybrid was also analyzed in depth to ensure its suitability with the characteristics of the original data, especially for the minority Worms attack class. The evaluation was carried out through t-SNE (t-Distributed Stochastic Neighbor Embedding) visualization, which is capable of projecting high dimensions into two dimensions so that distribution patterns can be clearly observed.

The visualization results show that the synthetic data points (marked in red) are widely distributed and form a manifold structure similar to the distribution pattern of the original data in the feature space. Although the amount of original Worms data is very limited—only one original data point (marked in blue) is visible in the graph—the synthetic data generated successfully builds a dense and organized distribution, illustrating the model's understanding of the topology of the minority feature space. This phenomenon confirms that the SDAE-GAN hybrid does not merely perform crude duplication, but is capable of replicating the complex intrinsic variations of this class of attacks.

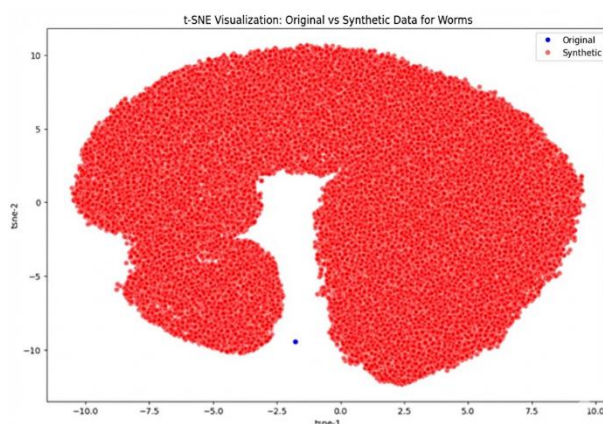


Figure 5. T-SNE Visualization of Original and Synthetic Data for The "*Worms*" Class

This finding contrasts with traditional oversampling techniques such as SMOTE, which often place synthetic points in the interpolated space between clusters and potentially introduce noise. In contrast, the SDAE-GAN

hybrid produces synthetic samples that reside on a realistic manifold structure and integrate naturally into the representation space of the original data. This shows that the pre-training stage using SDAE provides a strong representation foundation, allowing the GAN training process to be directed towards learning a more accurate and stable data distribution.

Thus, the visualization results using t-SNE show that the distribution of synthetic data in the Worms class is of good quality and able to represent the characteristics of the original data more realistically. The pattern of point distribution in low-dimensional space shows that the augmented data is not only randomly distributed, but follows a distribution structure that corresponds to the actual attack pattern. This finding provides empirical evidence that the SDAE-GAN hybrid augmentation approach successfully overcomes the problem of sample scarcity in rare attack classes. In addition, the model's ability to generate richly varied synthetic data helps the classification algorithm learn anomaly patterns more deeply and comprehensively. Therefore, this method has been proven to contribute significantly to improving detection performance in network attack classification systems.

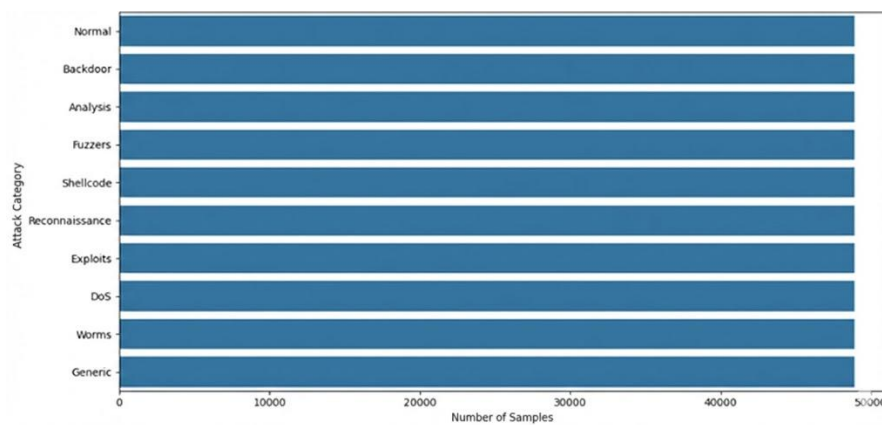


Figure 6. Class Distribution After Augmentation

Classification Model Performance

Testing using SVM on the augmented dataset showed a significant improvement in performance. Based on the experimental results, the following performance was obtained:

Overall accuracy: 97%

F1-Score: 98.44%

Recall (attack class): 99.67%

The Recall value of 99.67% is the most important result because it shows the model's ability to detect almost all attacks. In the initial dataset, certain attacks such as Worms were generally detected less than 50% of the time or not detected at all due to the lack of samples and the dominance of the normal class. This jump in performance shows that augmentation through SDAE-GAN hybrid successfully reduces majority bias and allows SVM to form more inclusive decision boundaries for rare attack variations.

To confirm the superiority of this approach, the research results were compared with several recent studies (2020–2025 period) using the UNSW-NB15 dataset.

Table 2. Comparison Of Model Performance On The UNSW-NB15 Dataset

Method / Reference	Imbalance Handling Technique	Classificati on	Accur acy	F1- Scor e	Recall (Attack)	Description

Hybrid SDAE-GAN (Proposal)	Generative (Hybrid)	SVM (RBF)	97.00	98.4 4	99.67	Highest recall, best balance.
CNN- BiLSTM [13]	None / Standard	Softmax	-	95.4 5%	94.70	Lower recall without generative augmentation.
Random Forest [14]	Feature Selection	Random Forest	98.63	97.8 0	-	High accuracy but biased toward the majority (low FPR, potentially high FNR).
SMOTE + RF [15]	SMOTE	Random Forest	95.10	-	-	SMOTE improves accuracy but is susceptible to noise.
Standard GAN + CNN [16]	Standard GAN	CNN	93.27	93.9 9	-	Performance is lower than the hybrid method due to GAN instability.
CNN + GRU [17]	SMOTE	CNN-GRU	99.23	-	89.20	The accuracy is very high, but the attack recall (89%) is far below the proposed method.

The results show a common phenomenon, namely the "accuracy paradox." For example, the CNN+GRU method does indeed record an accuracy of 99.23%, but it is only capable of achieving an attack recall of 89.20%, meaning that around 10% of attacks are not detected. In the context of cybersecurity, failing to detect 1 in 10 attacks is a major operational risk.

In contrast, the proposed SDAE-GAN hybrid method maintains a very high Recall (99.67%) despite its slightly lower accuracy (97%). This advantage stems from the SDAE-GAN hybrid's ability to generate more challenging synthetic attacks (*hard examples*), allowing the SVM to learn much more precise classification boundary patterns. The Confusion Matrix results show that there are only 330 false negatives out of a total of

100,800 attack data cases, while the number of false positives is recorded at 2,858. This false positive rate can be considered a reasonable compromise in order to maintain a very small false negative rate. In a Security Operations Center (SOC) environment, a few false alarms are far more tolerable than the risk of an attack slipping through undetected [18].

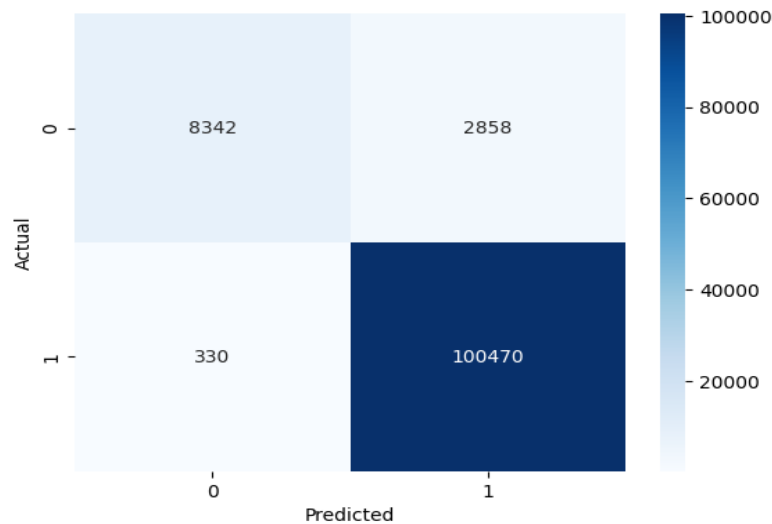


Figure 7. Confusion Matrix Of Model Testing On The Test Dataset

An additional evaluation was performed using the Receiver Operating Characteristic (ROC) curve to see the model's ability to distinguish between attack classes and normal classes at various decision thresholds. The ROC graph in the figure shows that the model testing curve is well above the baseline diagonal line (random classifier), with an Area Under Curve (AUC) of 98%.

This very high AUC value indicates that the model has a very strong discriminatory ability in recognizing attacks, even when the False Positive Rate (FPR) increases. The ROC curve appears to approach the upper left corner, indicating that the model is able to maintain a True Positive Rate (TPR) above 95% across most of the threshold range, before reaching a plateau close to TPR = 100%. This is consistent with the Recall result of 99.67% obtained in the classification evaluation.

Thus, the ROC curve analysis results further reinforce the previous findings that the SDAE-GAN approach not only successfully improves the proportion and balance of data distribution between classes, but also has a direct impact on increasing the sensitivity of the detection system. This improvement can be seen from the model's ability to maintain a high True Positive Rate (TPR) value while keeping the False Positive Rate (FPR) at a controlled level. This condition shows that the model is able to recognize attacks better without experiencing performance stability degradation. In the context of real-world implementation in intrusion detection systems in operational environments, the ability to maintain this balance is a very important factor. It is better for the system to generate a number of additional alerts than to allow dangerous attacks to slip through undetected. Therefore, the SDAE-GAN hybrid-based approach can be considered an effective and relevant solution for improving the reliability of IDS in the face of increasingly complex cyber threat scenarios.

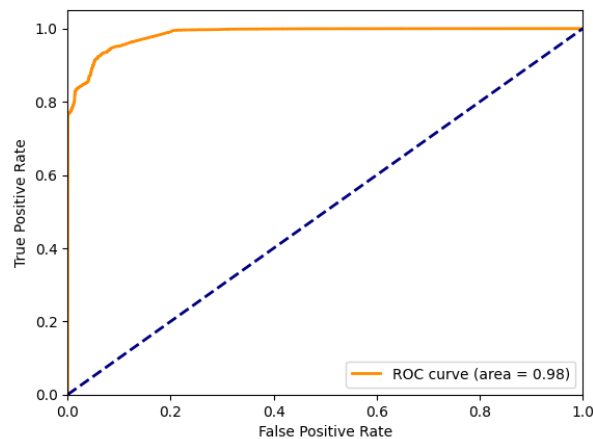


Figure 8. ROC Curve From Testing A Balanced Dataset

CONCLUSION

This study shows that the SDAE-GAN hybrid model is an effective approach in overcoming significant class imbalance problems in Intrusion Detection System (IDS) datasets. By utilizing the Stacked Denoising Autoencoder (SDAE) architecture as a generator in the Generative Adversarial Network (GAN) framework, this study successfully augmented the minority class, resulting in a more balanced training dataset. The Support Vector Machine (SVM) classification model trained using the augmented data showed excellent performance, particularly in the recall metric for detecting attacks, which is the main objective of IDS systems. An important contribution of this research is the proof that deep learning-based approaches are capable of producing high-quality synthetic data, thereby surpassing traditional augmentation methods such as SMOTE. This finding also confirms that with better data representation and balanced class distribution, conventional classification algorithms such as SVM are still capable of achieving near state-of-the-art performance. However, this research still leaves room for further development. Subsequent studies could explore more advanced GAN architectures, such as WGAN-GP, especially for tabular data. In addition, it would also be interesting to test the effectiveness of augmented datasets with various other deep learning-based classification models, such as CNN or LSTM, in order to identify the most optimal combination of methods.

REFERENCES

- [1] A. D. Vibhute, M. Khan, C. H. Patil, S. V. Gaikwad, A. V. Mane, and K. K. Patel, "Network anomaly detection and performance evaluation of Convolutional Neural Networks on UNSW-NB15 dataset," *Procedia Comput. Sci.*, vol. 235, pp. 2227–2236, 2024.
- [2] F. M. Anis, M. AlAbdullatif, S. Aljbli, and M. Hammoudeh, "A Survey on the Applications of Deep Learning in Network Intrusion Detection Systems to Enhance Network Security," *IEEE Access*, 2025.
- [3] M. Umer, M. Tahir, M. Sardaraz, M. Sharif, H. Elmannai, and A. D. Algarni, "Network intrusion detection model using wrapper based feature selection and multi head attention transformers," *Sci. Rep.*, vol. 15, no. 1, p. 28718, 2025.
- [4] S. Saiharish and D. Kumar, "Network intrusion detection through stacked machine learning models on UNSW-NB15 dataset," in *2024 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, IEEE, 2024, pp. 1–7.
- [5] J. J. Davis and A. J. Clark, "Data preprocessing for anomaly based network intrusion detection: A review," *Comput. Secur.*, vol. 30, no. 6–7, pp. 353–375, 2011.
- [6] E. M. Timofte *et al.*, "Federated learning for cybersecurity: A privacy-preserving approach," *Applied Sciences*, vol. 15, no. 12, p. 6878, 2025.

- [7] M. A. Umar, C. Zhanfang, and Y. Liu, "Network intrusion detection using wrapper-based decision tree for feature selection," in *Proceedings of the 2020 International Conference on Internet Computing for Science and Engineering*, 2020, pp. 5–13.
- [8] S. Debnath, "AI-driven cybersecurity for renewable energy systems: Detecting anomalies with energy-integrated defense data," *Int. J. Appl. Math. (Sofia)*, vol. 38, no. 5s, pp. 1002–1032, 2025.
- [9] W. Lim, K. S. C. Yong, B. T. Lau, and C. C. L. Tan, "Future of generative adversarial networks (GAN) for anomaly detection in network security: A review," *Comput. Secur.*, vol. 139, p. 103733, 2024.
- [10] M. Y. Jo and H. J. Kim, "A Comparative Study of Lightweight, Sparse Autoencoder-Based Classifiers for Edge Network Devices: An Efficiency Analysis of Feed-Forward and Deep Neural Networks," *Sensors*, vol. 25, no. 20, p. 6439, 2025.
- [11] Q. Fu and H. Wang, "A novel deep learning system with data augmentation for machine fault diagnosis from vibration signals," *Applied Sciences*, vol. 10, no. 17, p. 5765, 2020.
- [12] S. M. Kasongo and Y. Sun, "Performance analysis of intrusion detection systems using a feature selection method on the UNSW-NB15 dataset," *J. Big Data*, vol. 7, no. 1, p. 105, 2020.
- [13] H. Wang, Z. Cao, and B. Hong, "A network intrusion detection system based on convolutional neural network," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 6, pp. 7623–7637, 2020.
- [14] S. More, M. Idrissi, H. Mahmoud, and A. T. Asyhari, "Enhanced intrusion detection systems performance with UNSW-NB15 data analysis," *Algorithms*, vol. 17, no. 2, p. 64, 2024.
- [15] M. Sohail, M. Hamad, and A. Saeed, "NETWORK INTRUSION DETECTION USING THE UNSW-NB15 DATASET AND THE CONDITIONAL GANAUGMENTED CNN," in *Cryptography and Information Security Trends 2025*, Academy & Industry Research Collaboration, Sep. 2025, pp. 97–109. doi: 10.5121/csit.2025.151707.
- [16] G. Zhang, X. Wang, R. Li, Y. Song, J. He, and J. Lai, "Network intrusion detection based on conditional Wasserstein generative adversarial network and cost-sensitive stacked autoencoder," *IEEE access*, vol. 8, pp. 190431–190447, 2020.
- [17] L. Thomas, B. Iv, and A. Vakanski, "A Comparative Analysis of Machine Learning and Deep Learning Approaches for Network Intrusion Detection," 2025.
- [18] M. Lopez-Martin, A. Sanchez-Esguevillas, J. I. Arribas, and B. Carro, "Network intrusion detection based on extended RBF neural network with offline reinforcement learning," *IEEE Access*, vol. 9, pp. 153153–153170, 2021.