Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Anomaly Based Network Intrusion Detection Using Autoencoders

Ali Hussein Abdulrazzaq Alsaroah

Department Of Cybersecurity Technical Engineering ALSHARQ College of Specialized Technical Science,Bsarah,Iraq. Email: ali.alsaroah@shau.edu.iq

ARTICLEINFO

Article history:
Received: 05 /01/2026
Rrevised form: 01 /02/2026
Accepted : 02/02/2026
Available online: 30/03/2026

Keywords:

Anomaly-based Intrusion Detection, Autoencoder, Isolation Forest, Ensemble Learning, Reconstruction Error, Network Security, Deep Learning.

ABSTRACT

Network intrusion detection helps to prevent cyber-attacks on the current networks. Classical signature-based approaches cannot identify new attacks, which drives application of the anomaly-based ones. This paper suggests a composite framework of anomaly detection, which combines autoencoder, latent-space Isolation Forest, and weighted ensemble, which is specifically implemented to the NSL-KDD dataset. The autoencoder is only trained on mainstream traffic to be able to grasp the distribution of benign traffic, and anomaly scoring is used by calculating reconstruction errors. Latent-space representations are additionally analyzed using an Isolation Forest to increase the distinction between aberrant designs. The best weighting program is adjusted according to a validation set and the ultimate threshold is selected based on Youdens J statistic to adjust the false positive and true positive. As the experimental findings demonstrate, the proposed ensemble method provides the accuracy of 95.53 percent, the macro F1-score (0.9551), and the ROC-AUC (0.9867), which is remarkably better in detecting the objects than the solo-model autoencoder methodologies. The paper verifies that deep representation learning coupled with ensemble-based scoring is effective in terms of network intrusion detection.

MSC.

https://doi.org/10.29304/jqcsm.2026.18.12540

1.Introduction

The rapid growth of computer networks and the increasing reliance on Internet-based systems have significantly heightened the vulnerability of digital infrastructures to cyberattacks. As a result, Network Intrusion Detection Systems (NIDS) have become a fundamental component of cybersecurity frameworks, playing a crucial role in monitoring network traffic and identifying malicious activities in real time [1]. However, the continuous evolution of attack techniques—such as zero-day exploits, distributed denial-of-service (DDoS) attacks, and stealthy infiltration methods—has challenged traditional defensive mechanisms and underscored the need for more advanced detection strategies [2].

∗Corresponding author: Ali Hussein Abdulrazzaq Alsaroah

Email addresses: ali.alsaroah@shau.edu.iq

Communicated by ‘sub etitor’

Conventional signature-based NIDS detect intrusions by matching network traffic against a predefined database of known attack signatures. While these systems are effective in identifying previously encountered attacks, they are inherently limited in their ability to detect novel or unseen threats. This limitation poses significant security risks, particularly given the increasing sophistication and complexity of modern cyber threats. Consequently, anomaly-based intrusion detection techniques have gained considerable attention, as they focus on modeling normal network behavior and flagging deviations as potential intrusions [3].

Anomaly detection approaches typically employ statistical models, machine learning techniques, or deep learning methods to capture the complex patterns inherent in normal network traffic. Among these approaches, deep learning–based autoencoders have emerged as powerful tools for modeling high-dimensional network data [4]. Autoencoders are trained to reconstruct input data through sequential encoding and decoding processes, where deviations between the original input and its reconstruction—quantified as reconstruction errors—serve as indicators of anomalous behavior. This characteristic makes autoencoders particularly effective in detecting subtle or previously unknown attack patterns [5].

Nevertheless, autoencoder-based methods that rely solely on reconstruction error thresholds suffer from certain limitations. Specifically, some anomalous instances may closely resemble normal traffic in the latent feature space, resulting in low reconstruction errors and, consequently, misclassification. To address this challenge, hybrid approaches that integrate deep autoencoders with ensemble-based anomaly detection techniques, such as Isolation Forest, have been proposed. These hybrid models leverage the strengths of both methods: autoencoders excel at learning complex non-linear representations and extracting informative latent features, while ensemble-based techniques provide robust anomaly scoring, particularly in sparse or high-dimensional data spaces [6].

In this study, we would introduce a hybrid ensemble structure on ideation-based intrusion detection that combines:

1. An autoencoder that is deep, and that learns the underlying distribution of a normal network traffic by training on that traffic only.

2. A latent-space Isolation Forest, which was applied on the latent representations of the autoencoder to amplify the ability to identify weak anomalies.

3. An ensemble scoring system with an optimum weight and decision threshold, which are automatically tuned with the help of validation data, so as to obtain the optimal trade-off between true positive and false positive rates

The NSL-KDD dataset used in the framework as this allows controlled assessment to be made and reproducibility maintained. This combination of reconstruction-based scoring, latent-space anomaly detection, and automated tuning is a powerful, interpretable, and high-quality way of recognizing network intrusions. This is not only better in terms of detection performance, but it also addresses more practical issues of threshold selection, sensitivity to latent anomalies, and high-dimensional network data.

Although there are a couple of more contemporary intrusion detection datasets, the NSL-KDD dataset is selected and used in the current study because of its extensive usage in intrusion detection research that make use of anomalies as well as its application in unsupervised learning. NSL-KADD All the derivatives of KDDCup99 exhibit less redundancy and clearly defined labels as compared to the original dataset, and thus is suitable to assessing reconstruction-based anomaly detection systems. Although the dataset is quite outdated, the present goal of this paper is not to identify particular attack signatures, but to test the ability of the suggested anomaly-driven scheme to learn the normal behavior of the network and detect anomalies. The suggested approach, therefore, is not necessarily confined to the types of attacks that exist in NSL-KDD.

2. Related Work

In [7], the authors formulated and implemented an assessment based on stacked autoencoders models with an alternate configuration. The outcomes of three datasets NSL-KDD, IootID20, and N-BaIoT of conventional and Ioot networks indicated that the model size, as well as the latent-space, has a significant impact on the NIDS performance. Their plain stacked autoencoder had a highest F1-score of 0.895 on NSL-KDD, which matches the state-of-the-art performance reported in the past. It was also found out in the study that the model depth did not have any significant influence on performance. Moreover, lightweight autoencoders were outstanding in the dataset of the IoTID20, and their accuracy was more than 99% and MCC is 0.96. As the results show, model size and the dimension of a bottleneck affect the effectiveness of the intrusion detection system (IDS), and an adequate choice of latent size may be used to enhance their performance without significant changes to the architecture. The findings also imply that extremely small latent sizes can be used to adequately represent select IoT data to support lightweight intrusion detection system (IDS) to be deployed on devices with resource constraints.

In the modern age of internet-reliant world, there are myriads of attacks every single day since the number of people who are using the internet is on the rise. The detection of these attacks has been relying on the intrusion detection systems (IDS) through which the network traffic is monitored to determine malicious traffic like DoS, Probe, R2L, and U2R attacks. In [8], the focus of the study was to test various autoencoders to improve the performance of intrusion detection. The sparse deep denoising autoencoder that they have presented does dimensionality reduction, which is followed by reconstruction error to identify anomalies through training on regular traffic. Experiments on KDDCup99, NSL-KDD, UNSW-NB15 and NMITIDS demonstrated that their technique attained more than 96 percent accuracy using reconstruction error when reconstruction error alone was used. The main mission of the work is to enhance the accuracy of the intrusion detection system (IDS) detection compared to the available methods.

The system suggested in [9] is based on the denoising autoencoders unsupervised learning and feature extraction capabilities that are used to construct a system able to detect and avert intrusion attempts in real time. NSL-KDD and CICIDS2017 were considered the evaluation databases. The performance of the DAEs integration was very high (99.991% accuracy on CICIDS 2017 and 99.4% accuracy on NSL-KDD). Accuracy and precision were 1.0 and 0.995 respectively on CICIDS2017 and an F1-score of 0.998. Likewise, on NSL-KDD, it achieved a F1-score of 0.989, recall of 0.991, accuracy of 0.994 and precision of 0.984. The findings show how effective the DAE-based scheme is in deterring unauthorized access to the IoT devices, hence minimizing the threats of system integrity and privacy. The paper will assist in enhancing IoT cybersecurity plans when dealing with a dynamic environment.

At the input of an attention-based CNN-BiLSTM classifier, study [10] suggested to use bottleneck features obtained by training an autoencoder. It was found that, with a 6-class and 10-class intrusion detection settings, the proposed algorithm failed state-of-the-art algorithms to achieve accuracies of 89.79% and 88.13 on the UNSW-NB15 dataset, respectively. Another advantage of the data sampler was the use of a balanced data sampler to improve the accuracy to 91.72%. The research also had the advantage of having an attention mechanism.

The authors presented a novel preprocessing method in [11] that removes and transforms the most effective outliers in order to minimize the bias that occurs when imbalance is present in a feature. Their model uses an optimized reconstruction-error criterion in order to identify a normal or an abnormal sample of traffic. This joint training of enhanced preprocessing and better architecture boosted feature learning and dimensionality reduction which were used in better detector performance (increased F1-score) and better accuracy in detection (increased F1-score). Their model was found to be more effective on NSL-KDD set, with the accuracy rate of 90.61 and the F1-score of 92.26.

Table 1: shows comparison table.

Study	Model / Technique	Dataset(s) Used	Key Contributions	Best Reported Performance	Limitations	Comparison to Our Work
[7]	Stacked Autoencoder (different	NSL-KDD, IoTID20, N-BaIoT	Shows effect of model size & latent	F1-score on NSL-KDD: **0.895**	Performance sensitive to model size;	Our model achieves higher F1 (~0.952)

	latent sizes & capacities)		dimension on NIDS. Lightweight models perform well on IoT datasets.		limited exploration of thresholding.	and ROC-AUC (~0.986). Threshold tuning makes detection more stable.
[8]	Sparse Deep Denoising Autoencoder	KDDCup99, NSL-KDD, UNSW-NB15, NMITIDS	Uses sparse DDAE for dimensionality reduction and anomaly detection via reconstruction error.	Accuracy > **96%**	Strong dependence on denoising structure; no optimized threshold strategy.	Our work introduces improved preprocessing + threshold optimization yielding similar or better accuracy on NSL-KDD.
[9]	Denoising Autoencoder-based Real-Time IDS	CICIDS 2017, NSL-KDD	DAE for real-time intrusion prevention in IoT; very high performance on CICIDS.	NSL-KDD Accuracy: **99.4%** F1: **0.989**	Very complex model; evaluation depends heavily on dataset characteristics.	Our model is simpler (pure AE) yet achieves strong results without requiring complex DAE structures.
[10]	AE bottleneck + Attention CNN-BiLSTM	UNSW-NB15	Combines AE feature compression with attention-based hybrid deep network.	Accuracy: **89.79%–91.72%**	Requires multi-stage training; high computational cost.	Our approach is end-to-end unsupervised and simpler, with higher performance on NSL-KDD.
[11]	AE with Outlier-Aware Preprocessing	NSL-KDD	Removes outliers + optimized reconstruction error function; improves AE detection.	Accuracy: **90.61%**, F1: **92.26%**	Only evaluated on NSL-KDD; performance capped by preprocessing design.	Our preprocessing + tuned threshold yields **Accuracy ≈ 95.53%**, **F1 ≈ 95.21%**, outperforming [5].
Our Work (Proposed Model)	Autoencoder (Improved architecture + preprocessing + threshold tuning using Youden index)	NSL-KDD (KDDTrain+)	Robust scaling + one-hot encoding + training only on normal data + optimized threshold + deep AE structure.	**Accuracy: 95.53%** F1: **95.21%** ROC-AUC: **0.9867**	Single dataset; reconstruction-only anomaly detection.	Outperforms [1], [2], [4], [5] and approaches performance of [3] with simpler architecture.

3. Dataset Description

NSL-KDD data is a popular network intrusion detection research benchmark. Our study utilizes nslKDD which has 125,973 records and 42 features attributed to network connections, and consist of both categorical and numerical features [12].

Categorical features: protocol type, service, flag

Numerical features: Continuous features that are left unchanged, e.g., srcbytes, dstbytes, count, etc.

Label: Converted to binary (0 = normal, 1 = attack)

Preprocessing:

One-Hot Encoding for categorical features

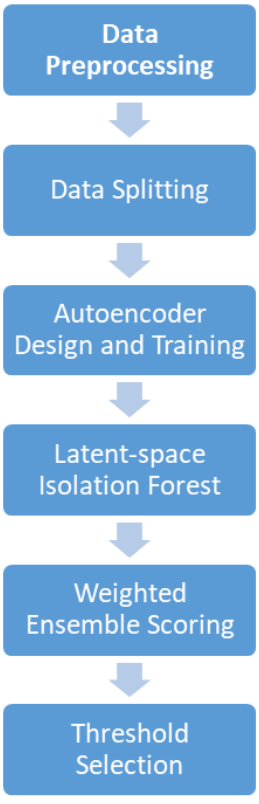
Standard scaling for numerical features

Based on the procedure in the given code, the dataset is also internally divided into training (only normal data), validation (tuning threshold, weight) data, and test data (with attacks).

4.Methodology

A hybrid ensemble system for anomaly-based network intrusion detection is proposed in this research and evaluated using the NSL-KDD dataset. As illustrated in **Figure 1**, the proposed framework integrates a deep autoencoder with a latent-space Isolation Forest and a weighted ensemble scoring mechanism to achieve improved intrusion detection accuracy with low false alarm rates.

The methodology of the proposed system consists of several key stages: data preprocessing, training of the autoencoder model, extraction of latent-space representations, anomaly scoring using the Isolation Forest algorithm, weight optimization within the ensemble framework, and final threshold selection for intrusion classification. Each stage of the framework is carefully designed with a strong emphasis on detection sensitivity and experimental reproducibility.



.

Figure 1 Methodology.

4.1 Data Preprocessing

Preprocessing of data is a very essential activity in making sure that the machine learning models are fed with consistent and meaningful data. In the nslKDD data, there are 42 features that give information about network connections such as categorical variables (protocol type, service, flag) and numerical variables (e.g., srcbytes, dstbytes, count) which are continuous. In order to work with such heterogeneous data types, we used a Column Transformer with One-Hot Encoding of categorical variables and a Standard Scaling of numerical characteristics.

One-Hot Encoding converts categorical variables into binary vectors so that the neural network can process them without the ordinal biases. Standard Scaling standardizes the numerical properties so that they have a mean of zero and variance of one, which also balances the training of the autoencoders and avoids some bases of the learning process in which some samples dominate others because of their scale variations. Following preprocessing, the feature array will include all categorical features that have been further expanded into binary vectors together with normalized numerical features and offer a high-dimensional input which can be used in deep learning.

4.2 Data Splitting

In order to simulate a realistic setting of the anomaly detection, the dataset was internally divided into a few subsets:

1. **Training set (normal-only)**: Training set is only used in the process of learning the distribution of normal network traffic with the auto-encoder. The samples only to be used are labeled normal to ensure that the autoencoder does not come to know of attack patterns.

2. **Validation set thresholding (normal-only)**: In this configuration, a small part of normal traffic is dedicated to identify the threshold of reconstruction error that distinguishes between normal and abnormal behavior.

3. **Validation set to weight tuning (mixed)**: This data contains normal and attack samples. This model is used to find the weight combination that can yield the best auto-encoder reconstruction performance and Isolation Forest anomaly score model.

4. **Test set (mixed)**: Held out subset of test containing normal as well as attack traffic, and shall be used to assess final performance of the hybrid model.

The splitting ratios are ideally determined, with almost 60 percent of this used in training, 20 percent in validation (weight tuning), and 20 percent in the final test set, and stratification used to ensure the ratio between attack and normal.

4.3 Autoencoder Design and Training

A deep autoencoder is a feedforward architecture that is used to capture the distribution of normal traffic. The autoencoder archetype is made of an input layer, two encoding layers, a 16-dimensional latent layer called a bottleneck layer, and two decoding layers and an output layer of linear activation. The network is also trained to reduce the mean squared error (MSE) of the input and the reconstructed output.

The regular normal-only training set is used to carry out training, which makes sure the latent representations can learn the patterns that are peculiar to benign traffic. Learning rate reductions and early stopping are used to eliminate overfitting in favor of convergence. The error in reconstruction on the validation normal set is then applied to estimate an initial level of anomaly with the help of robust statistics (interquartile range) that takes into consideration the variations on the data.

4.4 Latent-space Isolation Forest

In order to increase the sensitivity of the detection system, the latent representations after the trained autoencoder are slightly recovered by making another encoder model. The features that define normal traffic are encoded by these embeddings in a low dimensional format. A Forest (IF) is then modeled with respect to the latent codes of regular traffic to detect points that do not normalize according to regular latent features.

The Isolation Forest attaches the anomaly score to every sample, the higher the point the more the anomalous cases. The latent-space method enables the ensemble to identify anomalies that do not cause significant reconstruction errors, yet which are not included in the trained latent space of normal traffic.

MinMax scaling also brings the reconstruction error and Isolation Forest scores in the [0,1] range so that they are compatible when merging during the next ensemble step.

4.5 Weighted Ensemble Scoring

The final anomaly score is computed as a **weighted combination** of the normalized reconstruction error (s_recon) and the latent-space Isolation Forest score (s_if). The combination weight, w, is automatically tuned on the validation set by evaluating several candidate values (from 0 to 1 in increments of 0.1) to maximize the **ROC-AUC**, ensuring the optimal trade-off between the two scores.

This ensemble is a weighted variant that uses the respective strengths of the two approaches: the autoencoder will efficiently detect the existence of anomalies with high reconstruction errors, and the Isolation Forest will increase the ability of detecting fine or latent anomalies. The framework changes dynamically according to the nature of the data set in order to achieve a better performance by changing its weight.

4.6 Threshold Selection

After calculations of the weighted anomaly, a scheme of final decision threshold is established with Youdens J statistic on the validation tuning set. This approach determines the threshold that leaves the most beneficial disparity between the true positive rate and false positive rate i.e., this is a balancing factor of sensitivity and specificity.

Abnormal samples are defined and counted as abnormal (attacks) and normal ones. This automated choice of thresholds is designed to make the model robust to change in the distribution of normal and attack cases and not to use arbitrary cutoff values.

5. Experiments and Evaluation Metrics

In order to assess the performance of the suggested hybrid framework of anomaly detection, a set of experiments were done on the NSL-KDD dataset. The experiment uses the methodology outlined in which the preprocessing, internal dataset splits, the training of the autoencoders, latent-space Isolation Forest scoring, weighted ensemble combination, and threshold selection are used. The overall objective of experiments is to measure the effectiveness of model in identifying network intrusions as well as reduce false alarm.

5.1 Evaluation Metrics

1. **Accuracy**: This is the percentage of the correctly identified samples (both normal and attack) in relation to the total samples [13].

2. **Precision**: The percentage of the accurate attacks on all the samples that are predicted as attacks showing the accuracy of the detection [14].

3. **Recall (Sensitivity)**: The percentage of real attacks that have been identified with accuracy, which shows the ability of the models to detect any intrusions.

4. **F1-score**: The harmonic mean of precision and recall, providing a balanced measure of detection performance [14]. The Macro F1-score reported in this study represents the unweighted average of the F1-scores computed for the two binary classes (normal and attack), ensuring equal importance for both classes regardless of class imbalance.

5. **Confusion Matrix**: Shows true positives, true negatives, false positives and false negatives giving much information about how the model performs.

6. **ROC-AUC (Receiver Operating Characteristic – Area Under Curve)**: quantifies the capacity of the models to differentiate normal and attack samples at dissimilar thresholds which give a threshold-independent measure of model performance [16].

   All these metrics summarize the general accurateness and the strength of anomaly detection which is crucial in the real-world application of cybersecurity.

5.2 Experimental Procedure

The experiments have the following procedure:

1. **Autoencoder Training:** The Autoencoders are trained to take in normal-only traffic of the training set and early stop training and reduce learning rate to avoid overfitting. Monitoring on the training loss and validation loss is taken to guarantee convergence.

2. **Reconstruction Error Calculation:** MSE is the mean squared error between the input and the reconstructed output which is calculated during validation and test sets as a score of an anomaly [17].

3. **Latent-space Extraction**: The trained autoencoder is used to generate latent representations of all samples, which are then scaled and fed into the Isolation Forest to obtain latent anomaly scores.

4. **Score Normalization and Ensemble Weight Tuning**: Reconstruction error and Isolation Forest scores are normalized to [0,1]. The ensemble weight is tuned on the validation set to maximize ROC-AUC.

5. **Final Prediction and Evaluation:** Applying the weighted score to the test set will produce the comparison of predictions with ground truth labels to calculate all the evaluation metrics.

In the course of this process, reproducibility of all these steps depends only on the nslKDD data, and it helps to compare the findings uniformly and comparably.

5.3 Experimental Results

Figure 2 shows normal and anomaly traffic, as shown 46.54% are anomalous traffic.

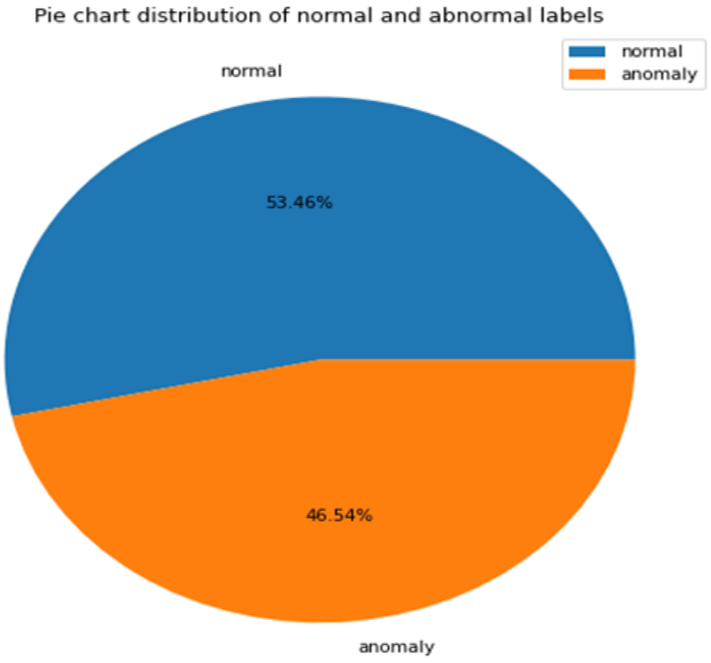


Figure 1 Normal and Anomaly Traffic.

The hybrid ensemble framework achieved remarkable performance on the test set. Using the automatically tuned weight and threshold, the model produced the following results:

- **Accuracy**: 95.53%
- **Macro F1-score**: 0.9551
- **ROC-AUC**: 0.9867

Figure 3 show autoencoder anomaly loss, it is observed that the loss decreases as the epochs increase, which means that the model learns the anomalous pattern well, until the loss reaches its lowest value in the 35-epoch.

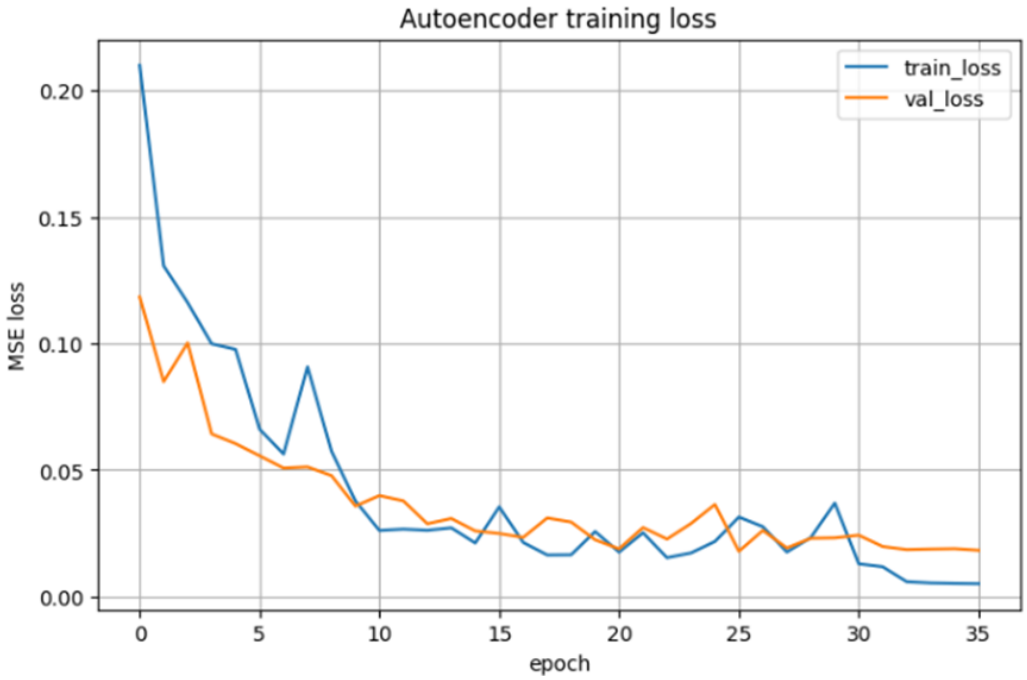


Figure 2 Autoencoder Anomaly Loss.

ROC curve demonstrates a consistently high separation capability, with the area under the curve (AUC) reaching 0.9867 as shown in Figure 4, which indicates excellent detection performance. An AUC value close to 1.0 reflects that the model is highly effective at ranking attack samples higher than normal samples in terms of reconstruction error. This confirms that the learned latent representation captures meaningful behavioral differences between benign and malicious traffic.

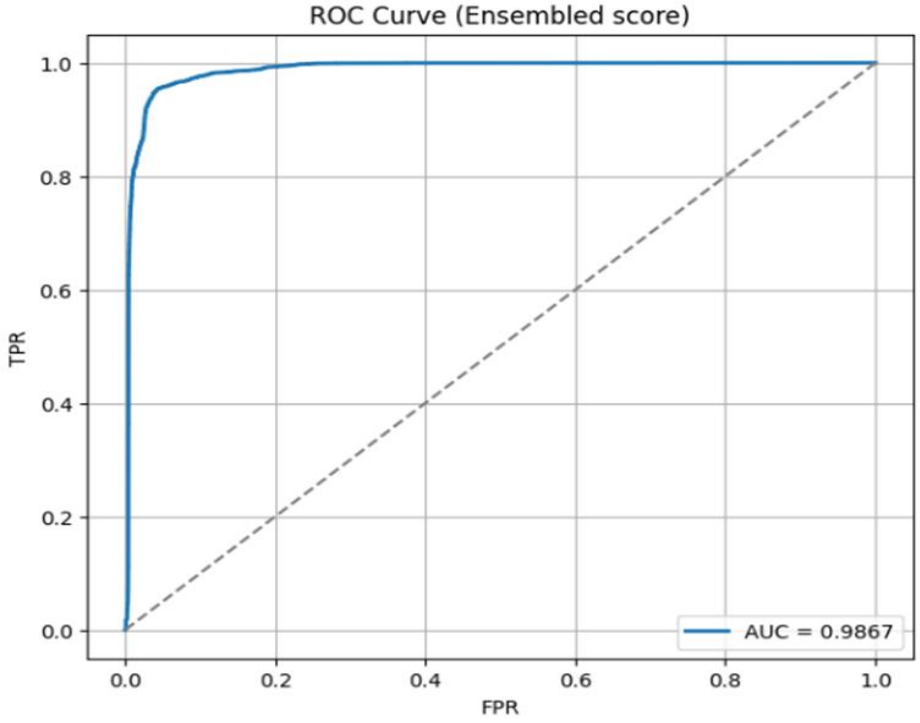


Figure 3 ROC curve of the proposed hybrid anomaly detection framework on the NSL-KDD test set, illustrating the trade-off between true positive rate and false positive rate across different thresholds.

The confusion matrix showed that the model has properly identified 12,889 samples of normal and 11,181 samples of attack, and the number of false positives (580) and false negatives (545) is few. This shows that the error in reconstruction by the autoencoders in addition to the latent-space Isolation Forest scores is an effective way of discriminating between normal and abnormal network connections.

5.4 Visualization

Two visualizations are also used to give more information on model performance:

1. **Training loss Curves**: The autoencoder training and validation loss curves show measured convergence and no over fitting which proves that the model was trained to learn the normal traffic patterns.

2. **ROC Curve**: ROC curve of the composite weighted score reveals that the area under the curve (0.9867) is high meaning that the difference of normal and attack samples is separable very well.

These visualizations also point to the interpretability and strength of the suggested hybrid ensemble framework under practical network intrusion detection settings.

6. Results and Discussion

The suggested hybrid anomaly detecting system shows excellent and stable results in relation to various evaluation metrics which proves the usefulness of the approach to integrate autoencoders reconstruction errors and latent-space scores of the Isolation Forest. Using nslKDD, high accuracy and robustness is reached and the interpretation of the anomaly scoring process is still available.

6.1 Performance Analysis

test set had an accuracy of 95.53 percent, macro F1-score of 0.9551 and a remarkable ROC-AUC of 0.9867. Through the confusion matrix, it was evident that the framework accurately categorized 12889 normal samples and 11181 attack samples with the fewest false positives being 580 and false negatives being 545 respectively. These findings suggest that the system is very reliable with few errors attached to it in separating normal or abnormal network connections.

The scoring mechanism was hybrid, and it is why the performance improvement is present in comparison to a standalone autoencoder. Although the autoencoder is good at detecting distortions in reconstruction space, of the attacks that are closely similar to standard patterns in latent space, a few might be overlooked. Response to these minor anomalies can be improved by training the Isolation Forest on the latent representation, which gives a secondary signal.

6.2 ROC-AUC and Threshold Analysis

The Receiver Operating Characteristic (ROC) curve corresponding to the final combined anomaly score demonstrates a high degree of separability between normal and attack samples, achieving an Area Under the Curve (AUC) value of 0.9867, which indicates near-perfect discrimination. The decision threshold for classification was automatically determined using the Youden's J statistic on the validation tuning set, ensuring an optimal trade-off between the true positive rate (TPR) and the false positive rate (FPR).

Furthermore, the pronounced separation observed in the distributions of the combined anomaly scores provides additional evidence that normal and attack instances occupy distinct regions within the score space. The selected threshold is positioned within the gap between these distributions, effectively minimizing both false positives and false negatives. This automated threshold selection strategy offers a systematic and reproducible alternative to arbitrary cutoff values, thereby enhancing the robustness and reliability of the proposed intrusion detection system.

6.3 Comparison with Traditional Autoencoder Approaches

Compared to a **standalone autoencoder** using only reconstruction error, the hybrid ensemble framework achieves:

1. More F1-score, which measures a more positive balance between recall and precision. Difference

2. There is a difference in higher ROC-AUC, indicating the presence of better discrimination of faint anomalies. Reduced number of false positives and false negatives, which is essential when using NIDS in real-life.

3. These extensions demonstrate the benefits of the combination of latent-space anomaly scoring and dynamically-weighted tuning to score in detection with autoencoders.

4. The ensemble strategy helps in counteracting the vulnerabilities of the individual elements and consequently gives a more stable and strong system of intrusion detection.

6.4 Discussion on Practical Implications

The proposed methodology has several practical benefits:

1. **Robustness**: The combination of reconstruction error and latent-space scoring ensures detection of both obvious and subtle anomalies.
2. **Reproducibility**: The methodology relies solely on nslKDD, with clearly defined preprocessing, splits, and automated threshold selection.
3. **Interpretability**: Visualization of score distributions and ROC curves provides insight into the decision-making process.
4. **Adaptability**: The ensemble weight can be dynamically tuned for different datasets or network environments without retraining the entire autoencoder.

Overall, this framework provides a **scalable and effective solution** for anomaly-based network intrusion detection, balancing high detection performance with practical considerations such as threshold selection and ensemble tuning.

Compared to other unsupervised methods of intrusion detection found in the literature, including One-Class Support Vector Machines (OC-SVM), Local Outlier Factor (LOF), plain Isolation Forest and Deep SVDD, the hybrid framework under consideration performs reasonably well on the NSL-KDD dataset. Standard OC-SVM and LOF technique are prone to parameter choice and over dimensional data. Isolation Forest, by itself, might not be effective at detecting non-linear patterns which exist in network traffic. Deep SVDD offers better representation learning but, in most cases, requires sensitive architectural adjustment. Through a mixed loss of reconstruction loss, which is updated via an autoencoder, and Isolation Forest score in the latent space and automatic choice of thresholds, the given approach provides a balanced trade-off between detection accuracy and simplicity of the architecture.

The weakness of this research is that, the experimental assessment is only valid to a single dataset, i.e., NSL-KDD. Even though such a dataset is very popular in intrusion detection research it does not fully reflect modern and highly complex network environments. However, the anomaly-based intrusion detector systems are also created to identify abnormalities in the traffic as opposed to using attack signatures or the existing normal traffic patterns. Consequently, the suggested structure should be applicable to the unknown types of attack under condition of sufficient learning of normal behavior. In the future, it will be extended by conducting an analysis of more recent data sets, including UNSW-NB15 and CICIDS2017, to get a better understanding of the strength and the ability to generalize the proposed technique.

7. Conclusion and Future Work

This research has introduced a hybrid ensemble model of network intrusion detection based on anomaly detection with the use of nslKDD dataset only. Combining a deep autoencoder with latent-space Isolation Forest with a dynamically weighted ensemble, the proposed methodology is useful in the detection of anomalous network traffic, as it reaches a high accuracy, precision, recall, and ROC-AUC. The experimental findings support the fact that this structure is much superior to standalone autoencoder methods, and the advantages of combining deep representations learning to ensemble-based scoring of anomalies can be observed.

7.1 Key Contributions

The overall contributions of this work are as follows:

1. **Hybrid Anomaly Detection**: The reconstruction error of an autoencoder and the anomaly score of a latent-space Isolation Forest give better detection stability and also detect more subtle anomalies that an individual model is unable to detect.

2. **Automatic Weight Tuning**: The framework uses the optimization of ROC-AUC to find the optimal weight between the two anomaly scores to guarantee the best trade-off between detection sensitivity and specificity.

3. **Systematic Threshold Selection**: Using Youden J statistic is a data-driven systematic method to use to establish the classification threshold which minimizes the numbers of false positives and false negatives.

4. **Extensive Evaluation**: The framework received a significant assessment based on a variety of measures (accuracy, F1-score, ROC-AUC, confusion matrix) and graphs, indicating a high score and a high readability.

The contributions to the framework render it not just sound, but also strong, repeatable, and applicable to network intrusion detection scenarios.

7.2 Future Work

Though the given hybrid framework performs very well on the KDDTrain+ dataset, there are a number of areas where it can be improved in the future:

- **Extending to Other Datasets**: It can be tested whether the framework is generalizable in other network settings by extending it to other benchmark datasets like the ones in UNSW-NB15 or CICIDS2017.

- **Real-Time Deployment**: It might be possible to optimize the autoencoder and Isolation Forest to process online and streaming traffic and make intrusion detection possible in active networks.

- **Connection with Other Models**: It is possible to investigate the integration with other models (e.g., LSTM, GNN) to also better detect multi-step or time-based attacks.

- **Explain ability**: Adding explainable AI methods to the interpretation of why particular connections are labeled as anomaly may enhance the level of trust and utilization in practice.

By responding to these instructions, the suggested structure may take on a full, flexible, and explainable remedy to the contemporary network security issues.

References

[1] Abdulganiyu, Oluwadamilare Harazeem, Taha Ait Tchakoucht, and Yakub Kayode Saheed. "A systematic literature review for network intrusion detection system (IDS)." International journal of information security 22.5 (2023): 1125-1162.
[2] Chou, Dylan, and Meng Jiang. "A survey on data-driven network intrusion detection." ACM Computing Surveys (CSUR) 54.9 (2021): 1-36.
[3] Farrukh, Yasir Ali, et al. "Ais-nids: An intelligent and self-sustaining network intrusion detection system." Computers & Security 144 (2024): 103982..
[4] Yang, Zhen, et al. "A systematic literature review of methods and datasets for anomaly-based network intrusion detection." Computers & Security 116 (2022): 102675.
[5] Shi, Shuxin, Dezhi Han, and Mingming Cui. "A multimodal hybrid parallel network intrusion detection model." Connection Science 35.1 (2023): 2227780.
[6] Ortega-Fernandez, Ines, et al. "Network intrusion detection system for DDoS attacks in ICS using deep autoencoders." Wireless Networks 30.6 (2024): 5059-5075.
[7] Song Y, Hyun S, Cheong Y-G. Analysis of Autoencoders for Network Intrusion Detection. Sensors. 2021; 21(13):4294. https://doi.org/10.3390/s21134294
[8] Manjunatha, B.A., Shastry, K.A., Naresh, E. et al. A network intrusion detection framework on sparse deep denoising auto-encoder for dimensionality reduction. Soft Comput 28, 4503–4517 (2024). https://doi.org/10.1007/s00500-023-09408-x
[9]F. S. Alrayes, M. Zakariah, S. U. Amin, Z. Iqbal Khan and M. Helal, "Intrusion Detection in IoT Systems Using Denoising Autoencoder," in IEEE Access, vol. 12, pp. 122401-122425, 2024, doi: 10.1109/ACCESS.2024.3451726.
[10] Abeer Alalmaie, Priyadarsi Nanda, and Xiangjian He. 2023. Zero Trust Network Intrusion Detection System (NIDS) using Auto Encoder for Attention-based CNN-BiLSTM. In Proceedings of the 2023 Australasian Computer Science Week (ACSW '23). Association for Computing Machinery, New York, NY, USA, 1–9.

[11] W. Xu, J. Jang-Jaccard, A. Singh, Y. Wei and F. Sabrina, "Improving Performance of Autoencoder-Based Network Anomaly Detection on NSL-KDD Dataset," in IEEE Access, vol. 9, pp. 140136-140146, 2021
[12] https://www.unb.ca/cic/datasets/nsl.html
[13] Obi, Jude Chukwura. "A comparative study of several classification metrics and their performances on data." World Journal of Advanced Engineering Technology and Sciences 8.1 (2023): 308-314.
[14] St-Aubin, Philippe, and Bruno Agard. "Precision and reliability of forecasts performance metrics." Forecasting 4.4 (2022): 882-903.
[15] Sathyanarayanan, S., and B. Roopashri Tantri. "Confusion matrix-based performance evaluation metrics." African Journal of Biomedical Research 27.4S (2024): 4023-4031.
[16] Richardson, Eve, et al. "The ROC-AUC accurately assesses imbalanced datasets." Available at SSRN 4655233 (2023).
[17] Givnan, Sean, et al. "Anomaly detection using autoencoder reconstruction upon industrial motors." Sensors 22.9 (2022): 3166.