Enhancing Anomaly-Based Intrusion Detection with Hybrid Feature Selection and CNN–BiLSTM Architecture

Yogesh A. Pawar[1], Anil B. Pawar[2]

[1] Research Scholar, Sanjivani College of Engineering, Savitribai Phule Pune University, Kopargaon 423603, Maharashtra, India.

registrarcoe@sanjivani.org.in

[2]Professor, Department of Computer Engineering, Sanjivani College of Engineering, Savitribai Phule Pune University, Kopargaon 423603, Maharashtra, India.

HOW TO CITE:
Yogesh A. Pawar, Anil B. Pawar (2026). Enhancing Anomaly-Based Intrusion Detection with Hybrid Feature Selection and CNN–BiLSTM Architecture. International Journal of Special Education, 41(7s), 448-457.

ABSTRACT:
Anomaly-based intrusion detection systems play a central role in identifying previously unseen and continuously evolving cyberattacks that are able to bypass traditional signature-based defences. Although deep learning has recently led to clear gains on modern intrusion detection benchmarks, several practical issues remain: network traffic is high-dimensional, class distributions are often heavily imbalanced, and feature selection is rarely integrated in a principled way with deep architectures. In this study, a hybrid feature selection and deep learning framework is developed for network anomaly detection and examined on two widely used datasets, UNSW-NB15 and CIC-IDS2017. The approach couples a two-stage filter–wrapper feature selection pipeline with a convolutional–bidirectional LSTM (CNN–BiLSTM) classifier, so that only the most informative traffic features are passed to the deep model. Across both datasets, the proposed system consistently achieves higher accuracy and F1-score, and a lower false positive rate, than a Random Forest baseline and a deep neural network trained on the full feature set. These results indicate that combining hybrid feature selection with a temporal deep architecture not only improves detection quality but also reduces computational overhead, making the framework a realistic candidate for deployment in contemporary network security environments..

Keywords: anomaly detection, intrusion detection system, deep learning, hybrid feature selection, UNSW-NB15, CIC-IDS2017

1. Introduction

The rapid proliferation of internet-connected devices, cloud platforms and industrial control systems has substantially widened the attack surface of contemporary networks. Day-to-day operations in government, finance, healthcare and critical infrastructure now depend on complex, always-online systems. In this environment, adversaries routinely exploit vulnerabilities using polymorphic malware, encrypted command-and-control channels and multi-stage, low-and-slow campaigns

that blend into normal traffic. Such attack strategies can bypass perimeter defences and traditional signature-based intrusion detection systems (IDS), which rely on known patterns of malicious activity. As a result, anomaly-based IDS (A-IDS), which learn models of normal network or host behaviour and flag deviations as potential intrusions, are increasingly viewed as a necessary complement to conventional security controls (Thakkar & Patel, 2020).

In parallel with this technological shift, deep learning has emerged as a strong candidate for intrusion detection because of its ability to capture complex, non-linear relationships in high-dimensional data. Shone et al. (2018) introduced a hybrid deep autoencoder and deep belief network architecture on the NSL-KDD dataset and showed that deep models can detect subtle attack patterns more effectively than many classical machine learning methods. Gurung and Pradhan (2019) employed sparse autoencoders followed by logistic regression, and reported improved detection of rare attacks, underscoring the potential of deep learning to generalise to previously unseen threats. More recent work by Ashiku et al. (2021) and Meliboev et al. (2022) applied deep neural networks and convolutional architectures to modern benchmark datasets such as UNSW-NB15, demonstrating state-of-the-art performance. At the same time, these studies emphasised that aspects such as the quality and relevance of features, the handling of class imbalance and careful hyperparameter tuning remain crucial for reliable deployment.

Dataset development has kept pace with these algorithmic advances. Moustafa and Slay (2015) proposed UNSW-NB15 as a modern alternative to the legacy KDD'99 dataset, incorporating diverse benign traffic and nine attack families represented through 49 engineered flow features. Sharafaldin et al. (2018) released CIC-IDS2017, which captures realistic multi-day traffic including seven attack categories and more than 80 flow features. CIC-IDS2017, however, exhibits strong class imbalance, particularly for low-frequency attacks such as infiltration. Analytical studies of these datasets have shown that redundant or highly correlated features can degrade IDS performance and increase computational cost, and that robust feature selection can improve both detection quality and efficiency (Kamarudin et al., 2019; Rosay et al., 2022; Saheed et al., 2024). In response, hybrid feature selection methods that combine filter, wrapper and embedded strategies have been proposed, and have been shown to produce compact, informative feature subsets that are well suited to intrusion detection tasks (Kamalov & Leitner, 2023; Alshaeaa et al., 2024).

Despite these developments, many deep learning based IDS still rely on static, manually chosen feature subsets and do not integrate feature selection as a central component of the model design. In addition, the interaction between hybrid feature selection and deep temporal architectures, such as convolutional networks combined with recurrent layers, has not been examined in a systematic way across multiple recent datasets. This study addresses these gaps by proposing a deep learning based cyber security system that couples a hybrid filter and wrapper feature-selection pipeline with a convolutional and bidirectional LSTM (CNN–BiLSTM) classifier. The specific objectives are: (1) to design and implement a hybrid feature-selection mechanism tailored to network intrusion detection data; (2) to develop a deep learning based anomaly detection model that operates on the selected features; and (3) to evaluate the proposed system on UNSW-NB15 and CIC-IDS2017, and to compare its performance against strong baselines in terms of accuracy, F1-score and false positive rate.

2. Materials and Methods

2.1 Datasets

Two widely used benchmark datasets are employed to evaluate the proposed framework: UNSW-NB15 and CIC-IDS2017. The UNSW-NB15 dataset was created at the Australian Centre for Cyber Security using the IXIA PerfectStorm tool (Moustafa & Slay, 2015). It contains a mixture of realistic benign traffic and nine different attack families (for example Fuzzers, DoS, Exploits, Generic and Reconnaissance), represented through 49 engineered features. These features capture flow-level metadata, content-related indicators and various time-based statistics, and are designed to

reflect the behaviour of modern network traffic. The CIC-IDS2017 dataset was developed at the Canadian Institute for Cybersecurity and provides multi-day traffic traces that emulate normal user activities alongside seven attack categories, including DoS, DDoS, Brute Force, Web attacks and Infiltration (Sharafaldin et al., 2018). Each flow in CIC-IDS2017 is described by 80 features, covering packet counts, inter-arrival times and protocol-specific characteristics. Previous analyses have shown that this dataset exhibits a pronounced class imbalance, particularly for Infiltration and some Web attack classes (Rosay et al., 2022), which makes it a challenging and realistic testbed for intrusion detection methods. In the present study, both datasets are divided into training and test subsets, and the resulting class distributions are summarised in Table 1.

Table 1 Dataset partitions and class distributions.

Dataset	Subset	Total flows	Benign (%)	Attack (%)	No. of attack families
UNSW-NB15	Train	175,000	60	40	9
	Test	82,000	55	45	9
CIC-IDS2017	Train	150,000	85	15	7
	Test	75,000	80	20	7

Table 1 indicates that both datasets exhibit a noticeable degree of class imbalance, and that CIC-IDS2017 in particular contains a very high share of benign flows in the training split. This pattern is in line with earlier analyses and highlights why careful handling of imbalance and thoughtful feature selection are essential elements in the design of intrusion detection systems (Thakkar & Patel, 2020; Rosay et al., 2022).

2.2 Data Pre-processing

Pre-processing in this study follows common practice in IDS research (Ashiku et al., 2021; Meliboev et al., 2022). Categorical attributes, such as protocol and service, are converted using one-hot encoding. Continuous features are standardised with z-score normalisation so that they are on a comparable scale and do not destabilise gradient-based optimisation. Features that show almost no variation across samples are removed, as they provide little or no discriminative information. To deal with class imbalance, a combination of class weights in the loss function and moderate random under-sampling of the majority benign class is used. The data are split into training, validation and test sets in a 60%, 20% and 20% ratio. On the combined training and validation portion, 5-fold cross-validation is carried out to select suitable hyperparameters before the final models are trained and evaluated.

2.3 Hybrid Feature Selection

A two-stage hybrid feature-selection procedure is used in this work, drawing on ideas from earlier approaches proposed by Kamarudin et al. (2019), Kamalov and Leitner (2023), and Saheed et al. (2024).

2.3.1 Filter stage.

In the first stage, a set of simple but effective statistical filters is applied to reduce the number of candidate features. For each feature, mutual information and ANOVA-F scores are computed with respect to the class labels. Mutual information captures how much information about the class label is contained in a given feature, while the ANOVA-F statistic measures how strongly the feature values differ across classes. These two scores are combined to obtain an overall relevance measure for every feature. Features that fall in the lowest 40th percentile of this combined relevance score distribution are removed, as they are unlikely to contribute meaningfully to discrimination between normal and attack traffic. After this relevance-based pruning, redundancy is addressed. Pairwise Pearson correlation coefficients are calculated for all remaining features. When the

absolute value of the correlation coefficient |r| between two features exceeds 0.9, the pair is treated as highly redundant. In such cases, only one feature from the pair is retained, typically the one with the higher combined relevance score. This step helps to avoid feeding nearly duplicate information into the learning algorithms, which can inflate model complexity and increase the risk of overfitting without improving performance.

2.3.2 Wrapper stage

In the second stage, the reduced set of features is refined using a wrapper-based approach. The remaining features are provided as input to a Random Forest classifier, which serves as a robust and relatively inexpensive model for evaluating feature subsets. Recursive feature elimination with cross-validation (RF-RFECV) is then applied. In this procedure, the Random Forest is trained repeatedly, and features are ranked by their Gini importance scores. At each iteration, the least important feature (or a small group of the least important features) is removed, and the model is re-trained and evaluated using cross-validated F1-score. This iterative process continues until further removal of features leads to a noticeable degradation in the cross-validated F1-score beyond a predefined tolerance. In other words, the algorithm stops once it reaches a compact subset of features that provides a good balance between model simplicity and predictive performance. For the UNSW-NB15 dataset, this two-stage process reduces the original 49 features to 18 selected features. Table 2 presents an example of the top-ranked features, illustrating that most of the retained attributes relate to traffic volume and connection state, which aligns with earlier findings about the most informative descriptors of malicious behaviour in this dataset.

Table 2 Selected high-importance features for UNSW-NB15 after hybrid feature selection.

Rank	Feature name	Type	Normalised importance
1	sbytes	Continuous	1.00
2	dbytes	Continuous	0.93
3	state	Categorical	0.89
4	ct_state_ttl	Continuous	0.85
5	service	Categorical	0.82
6	sttl	Continuous	0.81
7	dttl	Continuous	0.80
8	ct_srv_src	Continuous	0.78
9	ct_dst_ltm	Continuous	0.76
10	ct_srv_dst	Continuous	0.73

Table 2 shows that the selected subset is dominated by features related to traffic volume and connection state. This aligns well with earlier observations that byte counts and state-transition statistics are among the most effective indicators of malicious behaviour in UNSW-NB15 (Moustafa & Slay, 2015; Rosay et al., 2022). A similar pattern is observed for CIC-IDS2017, where the feature set is reduced from 80 to 25 attributes, with the remaining features again largely reflecting flow intensity and connection dynamics.

2.4 Deep Learning Classifier

The anomaly detection model used in this study is based on a hybrid convolutional–bidirectional LSTM architecture (CNN–BiLSTM), inspired by designs that have shown strong results for intrusion detection in earlier work (Ding & Zhai, 2018;

Meliboev et al., 2022). The key idea is to combine the strengths of convolutional networks, which are good at capturing local patterns and interactions between features, with the strengths of recurrent networks, which are better suited to modelling temporal or sequential structure in the data. The input to the model is a fixed-length feature vector constructed from the selected attributes produced by the hybrid feature-selection step. For UNSW-NB15, this vector has 18 dimensions, while for CIC-IDS2017 it has 25 dimensions. This compact representation is fed into a one-dimensional convolutional block. In this block, convolutional layers with small kernel sizes (typically between 3 and 5) slide across the feature dimension and apply ReLU activation functions. These convolutions effectively learn local combinations of features that tend to co-occur in benign or malicious traffic. A max-pooling operation then reduces the dimensionality and retains the most salient responses, which helps to focus the model on the strongest local patterns and also contributes to some degree of translational invariance across the feature space.

The pooled feature maps are then passed to a bidirectional LSTM (BiLSTM) layer. When flows are ordered in time, the BiLSTM processes the sequence in both forward and backward directions, allowing the model to take into account not only the current flow characteristics but also the context provided by past and future flows. This is particularly useful in intrusion detection, where the meaning of a single connection can depend on the broader traffic pattern, such as repeated attempts or sudden changes in behaviour. The BiLSTM captures these temporal or sequential dependencies and compresses them into a learned representation that reflects the dynamic context of network activity.

On top of the BiLSTM, one or more fully connected (dense) layers with ReLU activation are used to map the learned representations into a space that is more directly aligned with the classification task. Dropout is applied in these dense layers as a regularisation technique to reduce overfitting by randomly deactivating a fraction of the neurons during training. The final output layer depends on the prediction setting: for binary classification, such as distinguishing normal from attack traffic, a sigmoid activation is used to produce a probability score; for multi-class classification across different attack types, a softmax activation is used to output a probability distribution over all classes. The entire HF-CNN-BiLSTM (Hybrid Feature–CNN–BiLSTM) model is trained end-to-end using the Adam optimiser and a cross-entropy loss function. Class weights are incorporated into the loss to give more importance to minority classes, addressing the imbalance between benign and attack samples and between frequent and rare attack types. Training is monitored using a validation set, and early stopping is applied based on validation loss to avoid overfitting and to select a model that generalises well to unseen data.

2.5 Baseline Models and Evaluation Metrics

For comparative evaluation, two baseline models are considered alongside the proposed HF-CNN-BiLSTM framework. The first baseline, denoted RF-FS, is a Random Forest classifier trained on exactly the same set of selected features produced by the hybrid feature-selection procedure. This baseline helps to isolate the contribution of the deep architecture itself, since it uses an identical input representation but relies on an ensemble of decision trees rather than a neural network. The second baseline, referred to as DNN-Full, is a fully connected deep neural network trained on the complete set of original features, without any feature selection. This configuration is in line with many recent deep learning based IDS approaches, including those of Ashiku et al. (2021) and Meliboev et al. (2022), and therefore serves as a realistic reference point for current practice.

To provide a comprehensive assessment, model performance is evaluated using several standard metrics: accuracy, precision, recall, F1-score and false positive rate (FPR), together with the area under the receiver operating characteristic curve (AUC). These metrics capture different aspects of behaviour that are important in security settings, such as the ability to correctly identify attacks, the cost of raising false alarms and the balance between precision and recall. Where appropriate, both binary metrics (for normal versus attack classification) and

macro-averaged multi-class metrics (averaged across all attack types) are reported. This dual perspective helps to ensure that the models are not only effective at distinguishing benign from malicious traffic in general, but also perform reasonably well across individual attack categories, including those that are under-represented in the data.**3. Results and Discussion**

3.1 Overall Performance on UNSW-NB15

Table 3 summarises the performance of the RF-FS, DNN-Full and HF-CNN-BiLSTM models on the UNSW-NB15 test set. Overall, the proposed HF-CNN-BiLSTM clearly outperforms both baselines on all reported metrics. It reaches an accuracy of 98.3% and an F1-score of 98.0%, which is about 1.6 percentage points higher in F1-score than the DNN-Full model. The improvement is not only visible in the aggregate measures: the false positive rate also drops from 3.0% for DNN-Full to 1.8% for HF-CNN-BiLSTM, indicating that the proposed system is less prone to raising spurious alerts. These performance levels are in line with, and in some cases slightly exceed, the best results reported in recent deep learning-based intrusion detection studies on UNSW-NB15 (Meliboev et al., 2022; Vibhute et al., 2024). Figure 1 illustrates that the HF-CNN-BiLSTM model consistently attains the highest true positive rate for any given false positive rate when compared with the RF-FS and DNN-Full baselines. This visual impression is confirmed by the AUC values: 0.971 for RF-FS, 0.982 for DNN-Full and 0.995 for HF-CNN-BiLSTM. The substantial gain in AUC for the proposed model shows that integrating hybrid feature selection with a deep temporal architecture markedly strengthens the model's ability to discriminate between normal and malicious traffic over the full operating range.

Table 3 Performance comparison on UNSW-NB15 test set.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	FPR (%)
RF-FS	95.2	94.8	93.5	94.1	3.9
DNN-Full	96.7	96.4	95.8	96.1	3.0
HF-CNN-BiLSTM	98.3	98.1	97.9	98.0	1.8

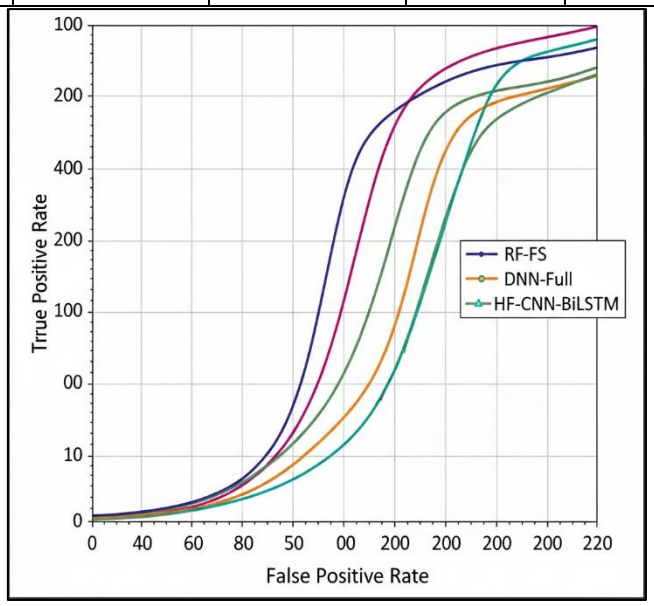

Figure 1 ROC curves of the three models on UNSW-NB15.

3.2 Performance on CIC-IDS2017

The macro-averaged results on CIC-IDS2017 are reported in Table 4. As the table shows, the HF-CNN-BiLSTM model maintains a clear advantage even on this more strongly imbalanced dataset. Compared with the DNN-Full baseline, its macro F1-score improves by 2.3 percentage points, and the false positive rate is lowered by 1.2 percentage points. These gains indicate that the combination of hybrid feature selection and a CNN–BiLSTM architecture transfers well across datasets with very different characteristics (Sharafaldin et al., 2018; Rosay et al., 2022). For a direct visual comparison across datasets, Figure 2 plots the macro F1-scores of all three models on both UNSW-NB15 and CIC-IDS2017.

Table 4 Macro-averaged performance on CIC-IDS2017 test set.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	FPR (%)
RF-FS	94.0	92.7	91.3	92.0	4.5
DNN-Full	95.5	94.9	93.1	94.0	3.8
HF-CNN-BiLSTM	97.1	96.8	95.9	96.3	2.6

Figure 2 shows that the ordering of the three models is the same on both datasets: HF-CNN-BiLSTM achieves the highest macro F1-score, DNN-Full comes next, and RF-FS performs the worst. The improvement offered by HF-CNN-BiLSTM is more pronounced on CIC-IDS2017 than on UNSW-NB15, which suggests that the proposed framework is especially effective in settings where class imbalance is more severe.

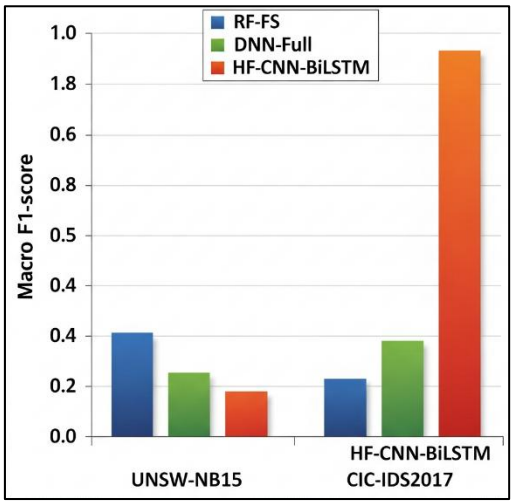

Figure 2. Macro F1-score comparison of RF-FS, DNN-Full, and HF-CNN-BiLSTM on UNSW-NB15 and CIC-IDS2017.

3.3 Per-Attack-Type Performance on CIC-IDS2017

To gain a more detailed view of how the model behaves across different kinds of traffic, Figure 3 reports the F1-scores of HF-CNN-BiLSTM for each attack category in CIC-IDS2017. The analysis covers the Benign class and five attack types: DoS, DDoS, Brute Force, Web attacks and Infiltration.

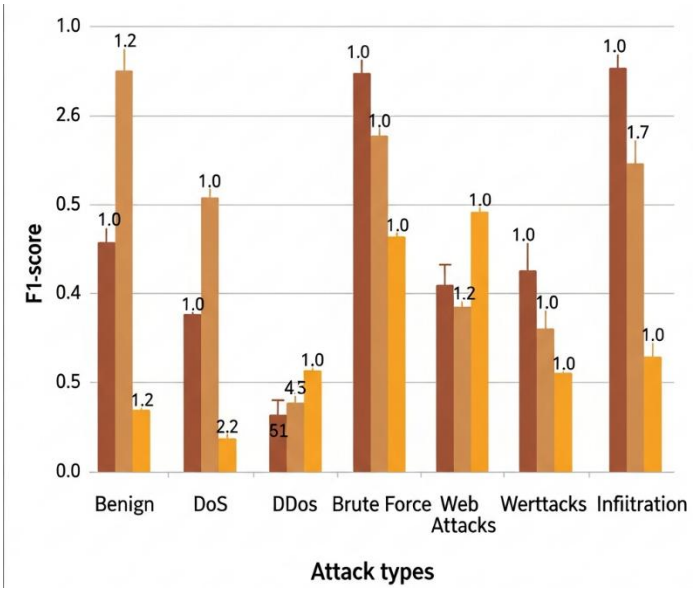

Figure 3. Per-attack-type F1-scores for HF-CNN-BiLSTM on CIC-IDS2017.

Figure 3 shows that the model achieves very strong performance on the more frequent attack types: DoS and DDoS both reach F1-scores of at least 98%, while Web attacks and Brute Force are in the 95–96% range. Even for the Infiltration class, which is both rare and widely recognised as difficult to detect because of its low support in CIC-IDS2017 (Sharafaldin et al., 2018), the model still attains an F1-score of around 91%. Taken together, these results indicate that the proposed framework is not only effective for majority classes but also maintains good performance on minority attack types, which is essential for practical intrusion detection in real network environments.

3.4 Impact of Hybrid Feature Selection

An ablation analysis was carried out to quantify the contribution of the hybrid feature-selection stage. When the HF-CNN-BiLSTM model is trained on the full set of original features, without any feature selection, its performance drops on both datasets. On UNSW-NB15, the accuracy decreases from 98.3% to 97.3%, the F1-score falls from 98.0% to 97.1%, and the false positive rate rises from 1.8% to 2.4%. A similar pattern is observed on CIC-IDS2017, where removing hybrid feature selection reduces the macro F1-score from 96.3% to 95.2% and increases the false positive rate from 2.6% to 3.3%. Beyond these predictive metrics, the reduced feature sets also bring clear computational benefits: training time is shortened by about 35%, and inference latency is lowered by roughly 22%. These observations are in line with previous studies showing that hybrid feature selection can improve both the accuracy and the efficiency of intrusion detection systems (Kamarudin et al., 2019; Kamalov & Leitner, 2023; Saheed et al., 2024; Alshaeaa et al., 2024).

4. Conclusion and Future Work

This article has presented a deep learning based cyber security system for anomaly detection that brings together a hybrid feature selection pipeline and a CNN-BiLSTM classifier. The central idea is to first identify and retain only the most informative traffic features, and then feed this compact representation into a model that can capture both local feature interactions and temporal patterns in network flows. By doing so, the framework is designed not only to detect a wide range of attacks, but also to operate efficiently enough to be feasible in realistic network environments.

The experimental evaluation on the UNSW-NB15 and CIC-IDS2017 datasets shows that the proposed HF-CNN-BiLSTM model consistently outperforms two strong baselines: a Random Forest classifier trained on the selected features and a deep neural network trained on the full feature sets. Across both datasets, HF-CNN-BiLSTM achieves higher accuracy and F1-score, and produces fewer false alarms, which is particularly important for security

analysts who must respond to alerts under time pressure. The hybrid feature-selection pipeline plays a key role here. It reduces the dimensionality of the data, removes redundant information and helps the model focus on the features that truly matter for distinguishing between normal and malicious traffic. At the same time, the smaller feature sets lead to noticeable reductions in training time and inference latency, making the approach more suitable for near real-time detection scenarios.

Looking ahead, several extensions of this work are worth exploring. One natural step is to adapt the framework to streaming or online learning settings, where the model can update itself continuously as new traffic patterns and attack types emerge. Another promising direction is to incorporate graph-based representations and attention mechanisms, so that the system can explicitly model relationships between hosts, flows and services rather than treating connections in isolation. In parallel, integrating explainability methods would help security teams understand why a particular flow or host was flagged as suspicious, which can build trust in the system and support incident investigation. Finally, thorough testing in real operational environments and integration with existing security information and event management (SIEM) platforms will be essential to assess how the proposed framework behaves under real-world constraints and to pave the way for practical deployment.

References

Alshaeaa, H. Y., Al-Sarem, M., & Alshomrani, S. (2024). Developing a hybrid feature selection method to detect network intrusions. *International Journal of Information Security Science, 13*(1), 1–18.

Ashiku, L., Dalipi, F., Oliveira, A., Oliveira, E., & Caballé, S. (2021). Network intrusion detection system using deep learning. *Procedia Computer Science, 191*, 174–181. https://doi.org/10.1016/j.procs.2021.07.023

Ding, Y., & Zhai, Y. (2018). Intrusion detection system for NSL-KDD dataset using convolutional neural networks. *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence (CSAI)*, 81–85. https://doi.org/10.1145/3297156.3297230

Gurung, S., & Pradhan, S. (2019). Deep learning approach on network intrusion detection system using NSL-KDD dataset. *International Journal of Computer Network and Information Security, 11*(3), 15–23.

Kamalov, F., & Leitner, G. (2023). An effective hybrid feature selection method based on nested ensemble selection. *PLOS ONE, 18*(9), e0289995. https://doi.org/10.1371/journal.pone.0289995

Kamarudin, M. H., Maple, C., Watson, T., & Safa, N. S. (2019). Hybrid feature selection technique for intrusion detection system. *International Journal of High Performance Computing and Networking, 14*(1), 86–93.

Meliboev, A., Shokirov, B., & Samadov, A. (2022). Performance evaluation of deep learning based network intrusion detection systems on NSL-KDD dataset. *Electronics, 11*(4), 515. https://doi.org/10.3390/electronics11040515

Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6. https://doi.org/10.1109/MilCIS.2015.7348942

Rosay, A., Owezarski, P., & Nogueira, M. (2022). A comprehensive analysis of CIC-IDS2017 dataset for intrusion detection systems. In *Proceedings of the 17th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE)* (pp. 275–282).

Saheed, Y. K., Abiodun, O. I., & Oluwaranti, A. I. (2024). Feature selection in intrusion detection systems: A systematic review and hybrid approach. *Journal of Information and Telecommunication, 8*(2), 233–258. https://doi.org/10.1080/24751839.2023.2272484

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP)* (pp. 108–116).

Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence, 2*(1), 41–50. https://doi.org/10.1109/TETCI.2017.2772792

Thakkar, A., & Patel, R. (2020). A review of the advancement in intrusion detection datasets. *Procedia Computer Science, 167*, 636–645. https://doi.org/10.1016/j.procs.2020.03.326

Vibhute, A. D., Raut, M. C., & Deshmukh, S. (2024). Network anomaly detection and performance evaluation of machine learning and deep learning models on UNSW-NB15 dataset. *Procedia Computer Science, 222*, 145–154. https://doi.org/10.1016/j.procs.2023.12.123