Volume 12, No.9, September 2025

Journal of Global Research in Mathematical Archives

ISSN 2320 – 5822

UGC Approved Journal

RESEARCH PAPER

Available online at http://www.jgrma.info

CYBERSECURITY ENHANCEMENT THROUGH DEEP LEARNING AND ANOMALY DETECTION FOR CYBER THREAT DETECTION

Dr. Amit Jain[1]

amit.jain@opju.ac.in

[1] Professor, Department of Computer Science and Engineering, OP Jindal University, Raigarh (C.G)

Abstract: The rapid advancement of digital technologies has made cybersecurity essential for safeguarding computer systems and networks. It involves techniques and algorithms to prevent unauthorized access, data breaches, and malicious activities while ensuring data security and privacy. A multi-layered defense on the host, application, network, and data levels is essential and continually enhanced to suit challenging cyber threats. With attack patterns getting increasingly advanced, cybersecurity has become a core element that supports trust and resilience in the digital age. This paper used the benchmark NSL-KDD data to consolidate an efficient cyber threat detection system. Data preparation includes filling in missing values, removing duplicates, cleaning up noise, one-hot encoding, normalizing, and using Random Forest to improve feature selection for a better model performance. Training and testing were conducted using deep learning and machine learning techniques such as ANNs, SVMs, DTs, and RNNs (Recurrent Neural Networks). At 96% accuracy, 92% precision, 93% recall, and 92% F1-score, RNN was the most successful model out of all that were evaluated. The findings indicate that combining feature selection and innovative deep learning methods remarkably enhances the detection accuracy, robustness and as such supports the increase in cybersecurity and securing against emerging cyber threats.

Keywords: Anomaly Detection, Cyber Threat Detection, Network Traffic Patterns, Intrusion Detection System (IDS), NSL-KDD dataset, Cybersecurity.

1 INTRODUCTION

The growing digitalization of industry and the increasing number of connected devices have significantly changed the face of contemporary business operations. All this digital transformation has enabled the activities of the IoT, cloud computing, and other digital platforms to grow exponentially, resulting in a complex and connected environment where continuously-developing data is exchanged [1][2][3]. As much as these developments have increased efficiency and innovation [4][5]This has led to new vulnerabilities, as industries are now more susceptible to cyber risks. The increasing dependence on digital infrastructure has also been accompanied by a sharp rise both in the sophistication and in the prevalence of cyberattacks, as cybercriminals take advantage of these vulnerabilities using continually more innovative practices [6][7][8].

Cybersecurity has become one of the pillars of the contemporary digital environment and has helped protect important information and intellectual property as well as critical infrastructure against malevolent activity [9]. The exponential growth of connected devices, our increasing dependence on cloud-based systems and IoT networks, and the resulting increase in the attack surface provide a perfect opportunity for sophisticated cyberattacks to exploit this weakness [10].

Conventional cybersecurity strategies effective in combating a particular set of threats would fail to keep the pace with the dynamism and complexity of contemporary cyberattacks [11]. This is illustrated by the fact that signature-based malware detection systems are unable to identify threats that have not been previously described, such as zero-day exploits or advanced persistent threats (APTs) [12]. Due to this, there is an increasing need to develop new solutions that can adjust and respond in a proactive manner to changes in cyber activity[13].

Since machine learning and big data analytics came along, supervised machine learning has been used as an alternative way to find hacking threats in real time [14][15]. ML is rapidly maturing into a technology that has the potential to revolutionize hacking techniques, particularly for the purpose of detecting threats in real-time. Traditional security methods might not be enough to stop cyberattacks because they are so advanced [16][17]. ML and big data analytics have been very important in making security systems able to find and stop threats in real time [18]. The speed and the amount of processed data provided by ML algorithms enable recognizing trends and anomaly and thus potentially recognizing the appearance of cyber threats. Machine learning can present a very viable alternative, and has superior features of locating, tracking and responding to threats with less inaccuracy and at higher speed [19].

1.1 Motivation and Contribution

The rationale of the research is the increasing demand in ensuring effective cybersecurity to mitigate the high number of advanced cyberattacks that are posing a global threat to systems, organizations, and individuals. The conventional security systems lack the ability to recognize new and sophisticated intrusion behavior and therefore, Intrusion Detection Systems (IDS) are necessary as an additional protection with respect to security. Despite this, many challenges exist with regard to the effective use of IDSs, including proper data preprocessing, selection of features, and application of learning schemes that can model complicated designs within the network traffic. By using the NSL-KDD dataset and orchestrating feature selection with the current models advanced technology like Recurrent Neural Networks (RNN), this study help improve the accuracy, reliability and efficiency in cyber threat detection, which eventually can heighten defensive cyber security during a most proactive way. The proposed research has the following major contributions to the subject of network security:

- Developed and tested intrusion detection models using the NSL-KDD dataset.
- Performed systematic preprocessing procedures, such as missing values treatment, removal of duplicates, noise elimination, one-hot encoding, and normalization in order to obtain high-quality input data.
- Used RF-based feature selection to expedite computations, remove duplication, and identify most important features.
- Demonstrated the applicability of the optimized framework of detection in real-life settings to enhance cybersecurity against dynamic malicious and sophisticated attacks.
- Constructed a RNN model that can be trained to recognize patterns in network traffic that occur sequentially and over time in order to effectively identify cyber threats.
- Ensured complete analysis by evaluating model performance using numerous measures, including recall, accuracy, precision, F1-score, and ROC.

1.2 Justification and Novelty

Traditional intrusion detection methods have its flaws, such as a propensity for false alarms, redundancy, and an inability to adapt to emerging cyber threats. These shortcomings provide the rationale for this study. Incorporating a range of advanced preprocessing approaches, feature selection with RF, and the RNN model, this study introduces a new framework that achieves the optimum balance of accuracy and performance. The innovation lies in the inclusion of optimized feature selection with the sequential learning capacity of RNNs, enabling the system to learn time patterns in network traffic that other models cannot recognize. The method not only saves on computational load but also leads to a powerful, more reliable, and scalable solution to current cybersecurity problem cases.

1.3 Structure of the paper

The structure of the paper is as follows: Section 2 surveys the relevant literature; Section 3 details the technique, including the dataset, preprocessing, and models; Section 4 presents the findings and examines them, while Section 5 provides conclusions and recommends more research.

2 LITERATURE REVIEW

Several important research studies in cyber threat detection to augment cybersecurity have been reviewed and assessed to inform and reinforce the formulation of this work.

Bavadiya et al. (2025) proposed a model that combines CNN-LSTM, RF, and autoencoders to enhance anomaly classification and detection. The Microsoft Malware Dataset, publicly available, was used as training and evaluation dataset. The AE-RF module trains on the normal behaviors of the system to detect abnormal behaviors, while the CNN-LSTM model learns the spatial and temporal characteristics of the malware to identify execution patterns. This article ensures that all data is pre-processed, specifically through feature extraction, PCA, and SMOTE, to balance the classes and provide optimal output per model. The experimental results indicate the AE-RF has a 0.91 AUC and the CNN-LSTM has 0.94, thus reflecting that the unsupervised and the supervised methods collectively exhibit high level of classification accuracy [20].

Khule, Motwani and Chauhan (2025) . This paper presents an Adaptive Threat Intelligence (ATI) framework that combines real-time model updates through incremental learning. The ATI framework does not require model retraining to accommodate new threat intelligence, as new intelligence is continuously integrated into the framework. The ATI framework uses a multilevel detection technique which includes an anomaly detection technique, behavior analysis and AI driven analytics technique. It also adheres to Zero-Trust security guidelines, which utilize continuous validation to prevent attacks from spreading. Experiments on actual datasets, such as threat logs from MITRE ATT&CK, demonstrate the efficacy of the proposed strategy, yielding noticeably higher detection rates than those of current techniques. ATI framework achieves an impressive 95.2% accuracy, 96.1% recall, and 94.8% F1-score, surpassing traditional models in detecting evolving APTs while reducing false positives [21].

Swetha and Merakapudi (2025) implement the two aforementioned unsupervised methods for detecting security abnormalities and new threats in real-time network environments. Results of this model evaluation are compared against standard supervised learning models, LR, LightGBM, and SGD Classifier using CICIoT23 dataset. The unsupervised hybrid model outperforms supervised models in accuracy measurements and maintains either equal or better recall and F1 score levels. The hybrid approach demonstrated

remarkable performance because it achieved 0.95 accuracy and 0.85 precision together with 0.94 recall and 0.89 F1 score, thus proving its effectiveness in detecting anomalies. The supervised models delivered accuracy measurements of approximately. Area under the ROC curve ratio between 0.87 to 0.89 and recall performance with a range of between 0.88 to 0.89 as well as F1 score with the highest measure of 0.89 [22].

Vemula et al. (2024) Their proposed solution is an anomaly detection system in the form of adversarial generative networks (GANs) that would detect unusual network usage patterns that are characteristic of attempted intrusion. In contrast to conventional methodologies, GAN can provide an adaptable, data-driven cybersecurity solution with an impressive purity of 96%, with DL models (90 -94%), unsupervised ML (92%) algorithms, hybrid genetic approaches (88%), and traditional NSL-KDD-based approaches (83%). This implies that GAN models can be used as a possible substitute since they deliver prodigious advancements in the network security malicious sounds [23].

Kalra et al. (2024) The proposed contribution presents a new methodology to the issue of identifying abnormal actions in traffic on the network based on the realization of CNNs. The model is trained and tested on a full range of normal and malicious network activities in the CICIDS2017 dataset. With CNN-based, an impressive accuracy of 93.64% was achieved, indicating a significant improvement over existing detection method. The findings, containing figures of merit and performance curves, precision, recall, and F-1 score, and confusion matrices perform a detailed evaluation of the model efficiency [24].

Kandhro et al. (2023) present a novel method for using DL to identify cyber-physical system vulnerabilities and assaults. Within the suggested structure, DL is contrasted with discriminative algorithms that rely on unsupervised learning. To identify cyber threats in IIC networks powered by the Internet of Things, this research proposes a generative adversarial network. The results show that all forms of assaults are identified with a performance improvement of about 95% to 97% when using an epoch value of 25 and a dropout value of 0.2. This improvement is evident in terms of efficiency, reliability, and accuracy. When it came to detecting BruteForceXXS, BruteForceWEB, DoS_Hulk_Attack, and DOS_LOIC_HTTP_Attack on the NSL-KDD, KDDCup99, and UNSW-NB15 datasets, the output of well-known state-of-the-art DL classifiers achieved the highest detection rate (HDR) and maximum true rate (TNR). Not only that, it protected private user and system information throughout training and testing [25].

Table 1 provides an overview of earlier research on cyber threat identification, highlighting novel models, datasets, significant findings, and the challenges encountered.

Table 1: Summary of Previous Studies on Anomaly Detection for Cyber Threat Using Deep Learning

Author	Proposed Work	Dataset	Key Findings	Challenges/recommendations
Bavadiya et al., (2025)	Hybrid AE-RF and CNN-LSTM model for anomaly detection & classification	Microsoft Malware Dataset	AE-RF achieved AUC 0.91; CNN-LSTM AUC 0.94; Combining unsupervised & supervised methods improves detection	Computational complexity; scalability to larger datasets
Khule, Motwani & Chauhan, (2025)	Adaptive Threat Intelligence (ATI) with incremental learning & Zero-Trust principles	MITRE ATT&CK threat logs	Achieved 95.2% accuracy, 96.1% recall, 94.8% F1; Outperforms traditional APT detection	Real-time adaptability requires robust infrastructure; may need optimization for large-scale enterprise deployment
Swetha & Merakapudi, (2025)	Unsupervised hybrid anomaly detection vs. supervised models	CICIoT23 dataset	With a precision of 0.85, recall of 0.94, and F1 of 0.89, the hybrid outperformed LR, LightGBM, and SGD.	Deployment in real-time IoT systems; scalability for high-speed traffic
Vemula et al., (2024)	GAN-based anomaly detection for intrusion detection	NSL-KDD dataset & others	Achieved ~96% purity; Outperforms DL (90–94%), ML (92%), hybrid (88%), classical (83%)	Training GANs is resource-intensive, potential instability in adversarial learning
Kalra et al., (2024)	CNN for detecting abnormal network traffic patterns	CICIDS2017 dataset	Accuracy 93.64%; Outperforms traditional models; Supports SDG 9 & SDG 16 goals	Needs testing on real-time systems; limited interpretability of CNNs
Kandhro et al., (2023)	GAN-based method for detecting IoT cyber threats in CPS	NSL-KDD, KDDCup99, UNSW-NB15	95–97% accuracy; Strong TNR & HDR; Maintained confidentiality & integrity	Requires fine-tuning GAN parameters; performance may vary across datasets

3 RESEARCH METHODOLOGY

This study's methodology begins with gathering the benchmark NSL-KDD dataset, which has been meticulously pre-processed to ensure reliable and high-quality data. The model's generalizability was enhanced during pre-processing by eliminating duplicate records, filtering out noisy data, and either removing or imputing missing variables. To ensure consistency and uniformity across

all features and to achieve zero mean and unit variance, Used One-Hot Encoding for categorical characteristics and the StandardScaler for continuous features to prepare the data for ML. Next, the RF algorithm was used to choose the features. This approach ranks features according to their importance, allowing for the identification of the most valuable ones for intrusion detection. The data was split in half for the purpose of training and testing the models; one half was utilised for training and parameter tuning, and the other half was used for testing and validation. The main model, an RNN, was finally introduced. Its ability to learn about sequential contexts and past event relationships makes it an ideal tool for understanding data patterns in time and sequential dependencies, both of which are crucial in determining whether network behavior is normal or abnormal. The proposed model was evaluated using the following popular metrics: accuracy, precision, recall, F1-score, and ROC curves, to ensure that the method effectively generated recommendations and classifications for cybersecurity improvement. Entire procedure of the methodology is shown in Figure 1.

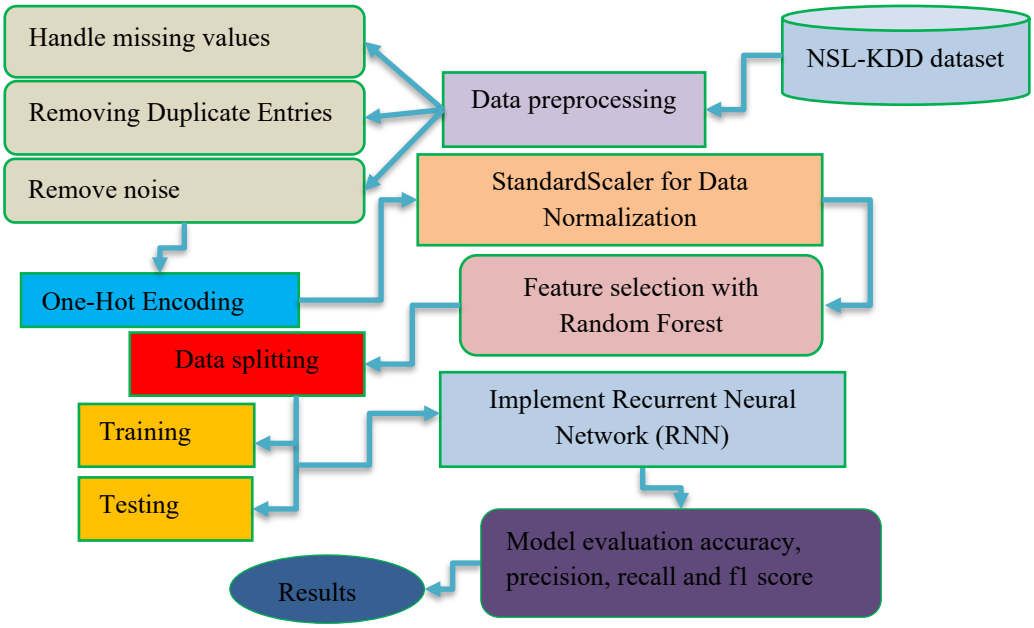

Figure 1: Proposed flowchart for Cyber Threat Detection

Cyber threat detection flowchart for improved cybersecurity, with thorough explanations of each step, is provided below:

3.1 Data collection

The standard benchmark NSL-KDD datasets are utilized in this investigation. While both sets of data have the same number of columns, the NSL-KDD has 551+ records, while the old set has 489. Intrusion detection systems (IDS) were tested using this dataset to identify computer network attacks. Data visualizations such as bar plots and heatmaps were used to examine attack distribution, feature correlations etc., are given below:

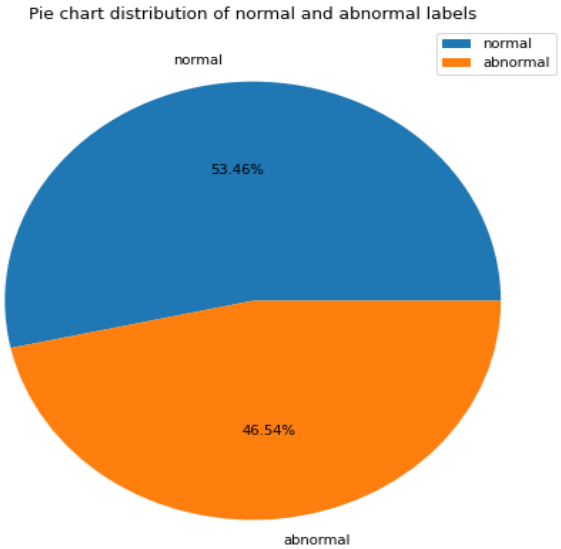

Figure 2: Binary classification

The distribution of normal and abnormal labels throughout the dataset is shown in Figure 2. Normal traffic is represented by 53.46% of the overall data, while aberrant traffic makes up 46.54%. This almost equal distribution guarantees that the dataset accurately portrays both benign and harmful actions, which is crucial for developing trustworthy intrusion detection models. The balance helps prevent bias towards either class, enabling the model to effectively learn and generalize in distinguishing between normal behavior and potential cyber threats.

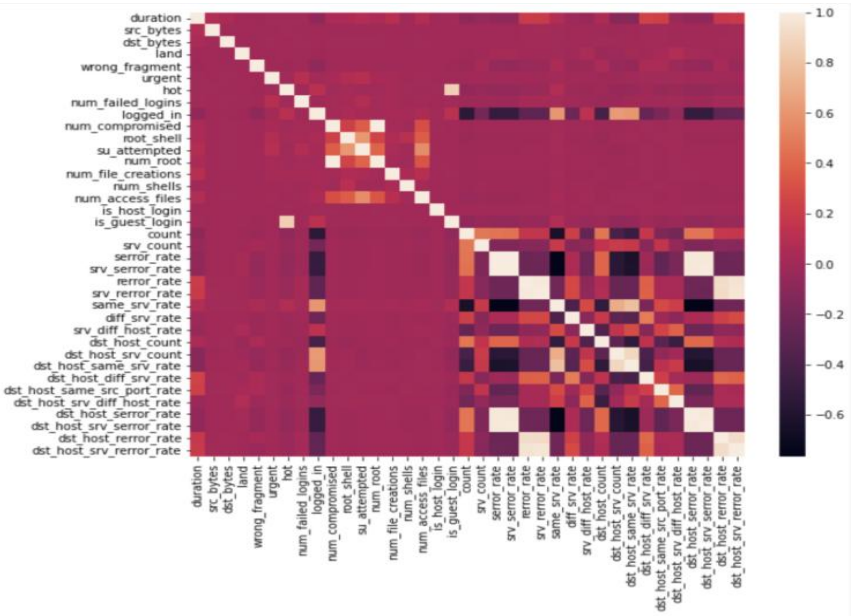

Figure 3: Heatmap for NSL-KDD Dataset

Figure 3 is a heatmap showing the dataset features' correlation matrix, where each column shows the correlation coefficient between two variables. The color scale ranges from dark shades indicating strong negative correlations to lighter shades representing strong positive correlations, with values close to zero shown in intermediate tones. The diagonal line consists of perfect correlations (value = 1) of each feature with itself. The heatmap reveals clusters of features that possess high interrelationships as well as remarkably independent features, yielding clues regarding redundancy and dependency, and possible multicollinearity across the dataset features.

3.2 Data pre-processing

The NSL-KDD data set was obtained, merged, and cleaned before being used in the data preparation process. The most important features were then selected. All of the duplicate entries, missing values, and noise in the dataset were removed during the pre-processing phase. Labelling and normalizing the data were also done. Below is a list of the preprocessing steps:

- **Handle missing values:** Missing values in a dataset can be handled using several techniques, including deletion, imputation (replacing with mean, median, mode, or other values), and more advanced methods.
- **Removing Duplicate Entries:** Select the data range wants to delete, then go to Excel's "Data" tab, and last, click the "Remove Duplicates" button.
- **Remove Noise:** Removing noisy data in machine learning is a crucial step in data preprocessing to improve model performance and generalization.

3.3 One-Hot Encoding for Data Encoding

Data encoding facilitates the efficient and effective handling of information by computers and other systems by transforming raw data into a predetermined format for storage, transfer, or processing. One-Hot Encoding makes use of binary variables for every category to convert categorical data to a numerical form. The numerical representation allows the model to learn the encoded attributes more effectively.

3.4 StandardScaler for Data Normalization

Given the different scales of each descriptor, the dataset was standardized using the StandardScaler () method to transform the data so that the mean of the resulting distribution is zero and the standard deviation is one. This transformation is achieved by subtracting the mean value of each observation and dividing by the standard deviation, as shown in Equation (1):

$$z = \frac{x-\mu}{\sigma} \tag{1}$$

In the dataset, z represents the feature's converted value, x depicts the descriptors' original values, μ denotes the feature's mean, and σ refers to its standard deviation.

3.5 Feature selection with Random Forest

The aim of feature selection is to construct a predictive machine learning model from a dataset by reducing the number of input characteristics and removing irrelevant or redundant ones. Then, the most relevant and helpful features are chosen. Due of its inherent capacity to priorities features, RF, an ensemble learning method, is frequently used by many organizations for feature selection. This method uses the decision tree structure of the forest to rank the attributes in order of significance when assessing the model's predictive potential.

3.6 Data Splitting

A training set and a testing set were created from the dataset so that efficiency could be measured. 80% of the time went into building the model and estimating its parameters, while the other 20% went into testing and evaluating its performance.

3.7 Proposed Recurrent Neural Network (RNN)

An RNN is a common kind of ANN in which the connections between the nodes produce a directed graph that contains information on the network's progress. The most effective usage of RNNs is in time series data, which is why the most recent and earlier data regulations yield the greatest results. A LSTM unit, which is used by RNNs, has an input gate, a forget gate, and an output gate within each memory cell [26]. To manage enormous data sets and address the gradient explosion issue, RNNs are used. The computation of the current state made use of this memory property [27][28]. Consequently, RNN units take in two values: the current input value and the output from before. One way to describe how RNN units' function is as (2).

$$y(t) = f\big(x(t) + y(t-1)\big) \tag{2}$$

$x(t)$ represents the input at the moment, $y(t)$ stands for the output at the moment, and $y(t-1)$ represents the output at the time before. The symbol $f(x)$ represents the function that models this relation. When it comes to RNN, the function $f(x)$ is a straightforward hyperbolic tangent function that causes nonlinearity due to its unstable gradient tendency. When activation functions like the rectified linear unit function are not saturated, gradients can get arbitrarily big. As demonstrated in figure 4, this model consisted of two recurrent layers and a single dense layer.

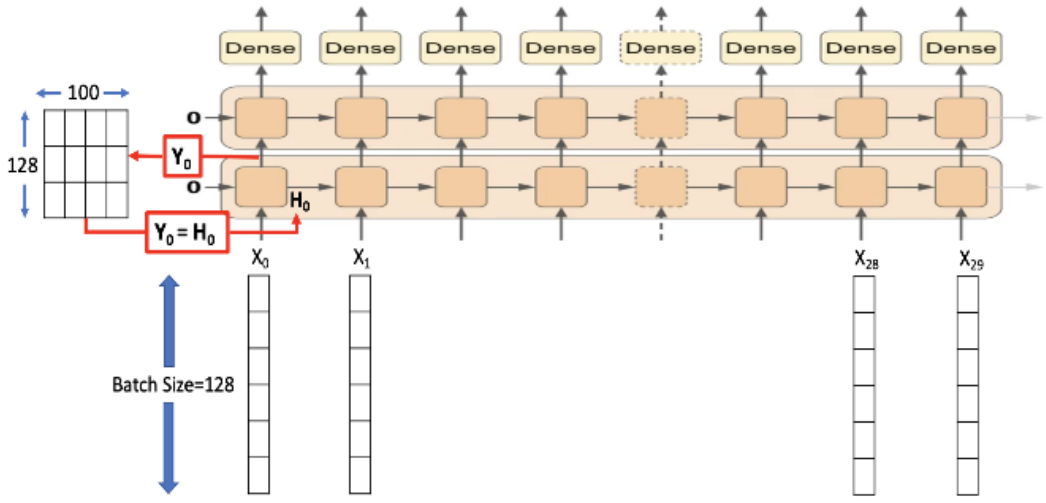

Figure 4: RNN model structure

A batch size of 128 is acceptable for the recurrent layers. There are 100 units in each recurrent layer, and each layer contains 30 time steps. By taking the results from the recurrent layers, the dense layer may forecast how many questions there with a single value. Within this design, RNN units are responsible for storing and passing on state information to subsequent iterations. Although stateful RNN models outperform their non-stateful counterparts on longer sequences, they are more time-consuming to train and can underperform on occasions due to strong correlations between batches.

3.8 Evaluation metrics

The proposed architecture was evaluated using a number of performance indicators. The results of TP, FP, TN, and FN were obtained by comparing the observed values with the expected outcomes from the trained models. These variables are used to construct the following performance metrics: F1-score, recall, accuracy, and precision:

Accuracy: The rate at which the trained model made predictions in comparison to the entire dataset (input samples) is measured by this statistic. The formula is (3)-

$$Accuracy = \frac{TP+TN}{TP+Fp+TN+FN} \quad (3)$$

Precision: Precision measures how well a model predicts positive occurrences relative to all positive occurrences. Precision indicates. How good the classifier is in predicting the positive classes is expressed as (4)-

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Recall: Positive event prediction accuracy is defined here as the proportion of true positives relative to the total number of false positives. In mathematical terms, it looks like (4)-

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

F1 score: It is a combination of the harmonic mean of precision and recall, that is, it helps to balance recall and precision. Its range is [0, 1]. Mathematically, it is given as (6)-

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (6)$$

Receiver Operating Characteristic Curve (ROC): The ROC plots, for a set of decision cut-off points, the ratio of successfully categorized cases to those that were wrongly classified. FPR is equal to 1-specificity, but TPR is often called sensitivity or recall.

4 RESULTS AND DISCUSSION

This section describes the experimental setup, along with the results of the training and testing steps for the proposed model. Machines that met the requirements for running the suggested algorithms included desktop PCs with Windows 11 Home 22H2, 16 GB of RAM, an AMD Ryzen 5-5500U CPU, and a Radeon 2.10 GHz graphics card. Anaconda Jupyter Notebooks powered by Python 3.8 were utilised for the development and execution of the ML models. Table II shows the training results of the proposed model on the NSL-KDD dataset. These results were derived from four important performance metrics: F1-score, recall, accuracy, and precision. Results from testing the suggested RNN model of cyber threat detection on the NSL-KDD database show that the model performs well across the board for the most important metrics. The model's 96% accuracy in evaluating it indicates a high level of capability to correctly categorise both known threats and normal traffic. Since the RNN is so good at avoiding false positives—it has a 92% accuracy rate—the threats it detects are incredibly trustworthy. The model's ability to capture the majority of real cyber threats, thereby reducing the likelihood of undetected attacks, is demonstrated by its 93% recall rate. In addition, the proposed RNN model has a robust F1-score of 92%, which means that it can out-perform the state-of-the-art in cybersecurity through accurate and trustworthy threat detection, without sacrificing any of the other important aspects of the system, such as recall and precision.

Table 2: Experiment Results of Proposed Models for Cyber Threat Detection to Increase Cybersecurity on NSL-KDD dataset

Performance matrix	Recurrent Neural Network (RNN) Model
Accuracy	96
Precision	92
Recall	93
F1-score	92

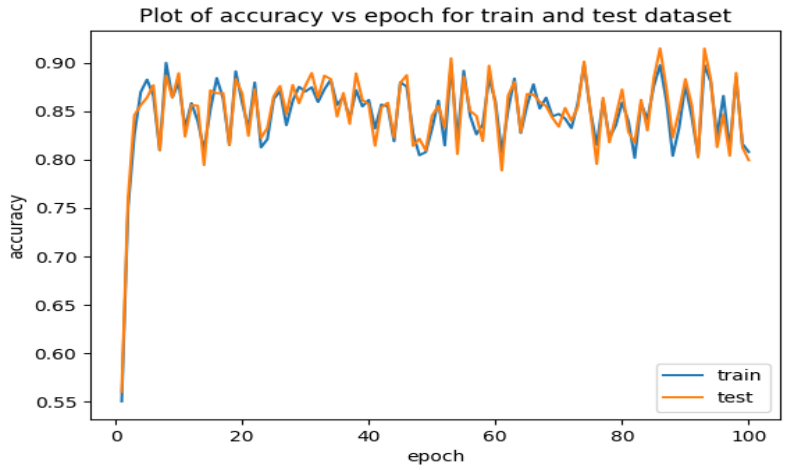

Figure 5: Accuracy curves for the RNN Model

Figure 5 demonstrates the accuracy of the proposed model during testing and training on the datasets of 100 epochs. Initially, the accuracy rate rises quickly in the first few of the epochs, increasing by approximately 0.30, and this is evidence of rapid learning in the earlier stages of training. Both the accuracy of training and testing the model reaches a plateau after the fourth training pass and fluctuates between 0.80 and 0.88; this is an indication that there is no notable overfitting or underfitting of the model. There is also tight convergence in the training and testing curve, which is an indication of the fact that the model has generalization ability and hence shows good and stable performance in previously unseen data.

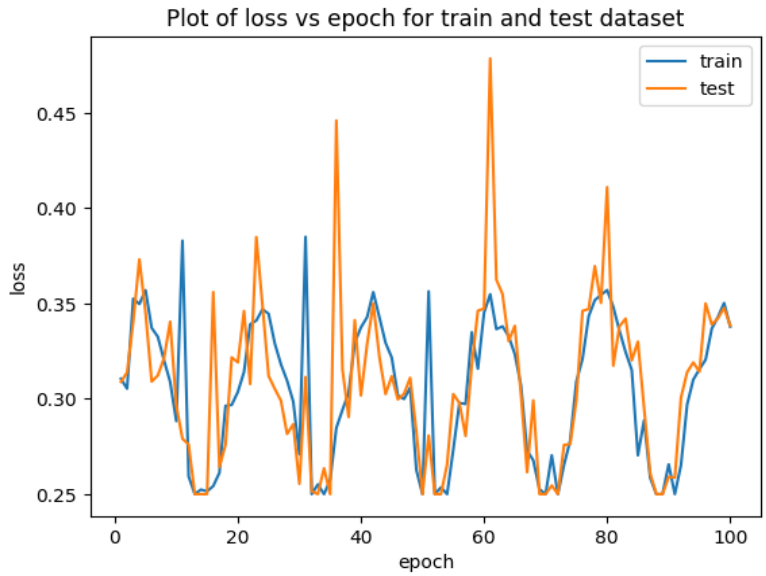

Figure 6: Loss curves for the RNN Model

Figure 6 indicates the loss change with each iteration in terms of epochs on the training and testing datasets within 100 epochs. Initially, the loss value is rather high, about 0.7, but it drops a lot in the first several epochs indicating that the model begins to learn successfully. As the training goes on, the loss reaches lower figures and stabilized at 0.25-0.35 indicating optimal model training. The loss curve of both the training and testing shows a similar monotony of growth with periodic peaks perhaps the interventions during learning or optimization. The close alignment of the two curves suggests that the model achieves good generalization, as it performs consistently on both the training and unseen testing datasets.

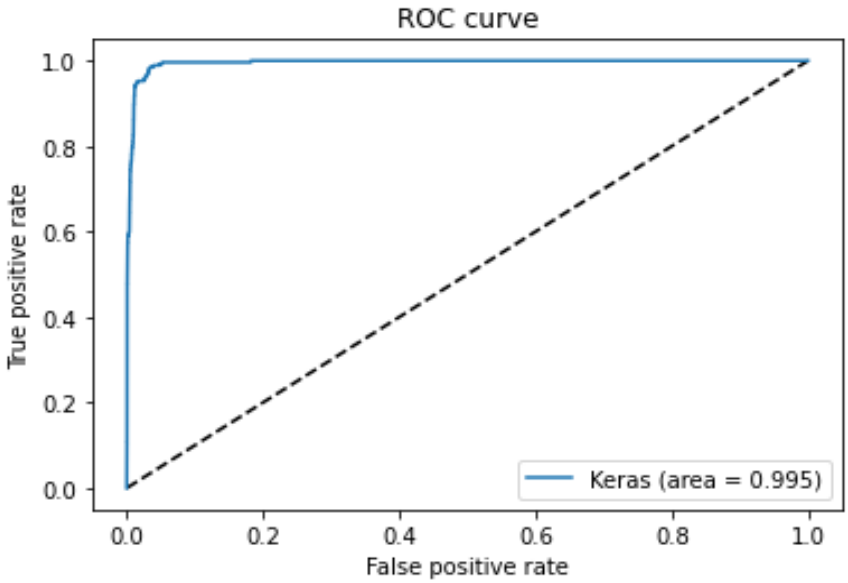

Figure 7: Receiver Operating Characteristic (ROC) curve for RNN

The suggested model's classification performance is shown in Figure 7, which is a plot of the sensitivity (actual positive rate) against the false positive rate (the ROC curve). Exceptional cross-class discrimination accuracy with low false positive rates is shown by the steeply rising curve in the top left corner. With an AUC of 0.995—very near to 1—the model works nearly perfectly and has excellent classification performance. The approach identifies cyber hazards quite accurately by taking sensitivity and specificity into account.

4.1 Comparative analysis

In order to confirm that the RNN model is functional, its accuracy is compared to other models that are already in use. Compared here. The NSL-KDD dataset was used to compare the prediction models' performance in detecting cyber threats (Table III). The most effective model was the ANN, which achieved an accuracy of 78% and a good level of judgement precision (96%). However, its lower recall (62.05%) and F1-score (75.57%) suggest that it did not detect all the relevant threats that were expected of it. The SVM achieved a somewhat better accuracy rate of 87.58% while maintaining steady performance in the other three metrics: recall (87.31%), precision (86.11%), and F1-score (87.07%). A 95.8 percent accuracy rate belies the DT model's weak generalizability and overfitting, as indicated by its low precision, recall, and F1-score values (around 55-56%). In comparison to the other models, the suggested RNN outperformed them all in terms of accuracy (96%), precision (92%), recall (93%), and F1-score (92%), demonstrating its potential to enhance cybersecurity by detecting threats more accurately and in a more balanced manner.

Table 3: Accuracy Comparison of different Predictive models of Cyber Threat Detection for increasing Cybersecurity using the NSL-KDD dataset

Models	Accuracy	Precision	Recall	F1-score
ANN[29]	78	96	62.05	75.57
SVM[30]	87.58	86.91	87.31	87.07
DT[31]	95.8	56.6	55.8	56.2
RNN	96	92	93	92

The main strength of the proposed RNN-based model is its high accuracy of 96%, which outperforms the effectiveness of other existing models on the NSL-KDD dataset related to cyber threat detection. This high accuracy indicates the ability of the model to correctly perform classification of the normal and malicious activities and thus minimizing mistakes in the identification of threats. With such high levels of accuracy, the proposed RNN provide a chance to introduce more trustworthy cybersecurity procedures, reducing the likelihood that possible attacks missed and increasing the efficiency of identification systems in general.

5 CONCLUSION AND FUTURE STUDY

Advanced persistent attacks, zero-day vulnerabilities, polymorphic malware, and other complex cyber threats are becoming more common and difficult for traditional malware detection approaches to identify. Using the NSL-KDD dataset and Random Forest analysis for systematic preprocessing and feature selection, this study demonstrates how to greatly improve the cyber threat identification problem. The method eliminates redundancy and trains only the most pertinent features resulting in improvements in computational costs and better overall model performance. Experimental results indicated that ANN attained 78%, SVM a high 87.58%, and DT 95.8%. It was observed that the RNN was able to predict the data most accurately with accuracy of 96%, proving that it can learn the temporal patterns of network traffic data efficiently. These findings confirm that optimized feature selection not only minimizes processing overhead but also strengthens predictive accuracy, enabling more reliable detection of intrusions. The study highlights the importance of implementing sophisticated cybersecurity models to address increasingly complex and intricate attack scenarios.

Improving the work's adaptability, scalability, and resilience against modern cyber-attacks could be achieved by testing state-of-the-art DL architectures on large-scale, real-time datasets. These designs include CNN-RNN hybrids, attention-based models, and transformers.

REFERENCES

[1] I. Adekuajo, O. Fakeyede, C. Udeh, and C. Daraojimba, “The Digital Evolution In Hospitality: A Global Review And Its Potential Transformative Impact On U.S. Tourism,” *Int. J. Appl. Res. Soc. Sci.*, vol. 5, pp. 440–462, 2023, doi: 10.51594/ijarss.v5i10.633.

[2] P. Mikalef and J. Krogstie, “Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities,” *Eur. J. Inf. Syst.*, 2020, doi: 10.1080/0960085X.2020.1740618.

[3] D. Patel, “Leveraging Blockchain and AI Framework for Enhancing Intrusion Prevention and Detection in Cybersecurity,” *Tech. Int. J. Eng. Res.*, vol. 10, no. 6, 2023, doi: 10.56975/tijer.v10i6.158517.

[4] V. Thangaraju, “Enhancing Web Application Performance and Security Using AI-Driven Anomaly Detection and Optimization Techniques,” *Int. Res. J. Innov. Eng. Technol.*, vol. 9, no. 3, 2025, doi: 10.47001/IRJIET/2025.903027.

[5] H. Kali, “The Future of HR Cybersecurity: AI-Enabled Anomaly Detection in Workday Security,” *Int. J. Recent Technol. Sci. Manag.*, vol. 8, no. 6, pp. 80–88, 2023.

[6] M. Yan, B. Chen, G. Feng, and S. Qin, “Federated Cooperation and Augmentation for Power Allocation in Decentralized Wireless Networks,” *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.2979323.

[7] V. K. Ayyam and V. Sudarsanan, “Zero-Touch Care Models: Evaluating Voice-Activated Health Assistants in Geriatric Palliative Care,” *Int. J. Sci. Res. Mod. Technol.*, vol. 4, no. 8, pp. 1–12, 2025, doi: 10.38124/ijsrmt.v4i8.726.

[8] W. Ahmad, A. Rasool, A. R. Javed, T. Baker, and Z. Jalil, “Cyber Security in IoT-Based Cloud Computing: A

Comprehensive Survey," *Electronics*, vol. 11, no. 1, p. 16, Dec. 2021, doi: 10.3390/electronics11010016.
[9] A. R. Bilipelli, "AI-Driven Intrusion Detection Systems for LargeScale Cybersecurity Networks Data Analysis: A Comparative Study," *TIJER*, vol. 11, no. 12, pp. 1–7, 2024.
[10] A. Whitmore, A. Agarwal, and L. Da Xu, "The Internet of Things—A survey of topics and trends," *Inf. Syst. Front.*, 2015, doi: 10.1007/s10796-014-9489-2.
[11] R. Q. Majumder, "A Review of Anomaly Identification in Finance Frauds Using Machine Learning Systems," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 101–110, Apr. 2025, doi: 10.48175/IJARSCT-25619.
[12] D. D. Rao, A. A. Waoo, M. P. Singh, P. K. Pareek, S. Kamal, and S. V. Pandit, "Strategizing IoT Network Layer Security Through Advanced Intrusion Detection Systems and AI-Driven Threat Analysis," *J. Intell. Syst. Internet Things*, vol. 12, no. 2, pp. 195–207, 2024, doi: 10.54216/JISIoT.120215.
[13] S. A. Bakhsh, M. A. Khan, F. Ahmed, M. S. Alshehri, H. Ali, and J. Ahmad, "Enhancing IoT Network Security Through Deep Learning-Powered Intrusion Detection System," *Internet of Things (Netherlands)*, vol. 24, 2023, doi: 10.1016/j.iot.2023.100936.
[14] C. S. Kruse, R. Goswamy, Y. Raval, and S. Marawi, "Challenges and opportunities of big data in health care: A systematic review," *JMIR Medical Informatics*. 2016. doi: 10.2196/medinform.5359.
[15] R. Patel, "Automated Threat Detection and Risk Mitigation for ICS (Industrial Control Systems) Employing Deep Learning in Cybersecurity Defence," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, pp. 584–591, Dec. 2023, doi: 10.14741/ijcet/v.13.6.11.
[16] M. Al-Hawawreh, N. Moustafa, and J. Slay, "A threat intelligence framework for protecting smart satellite-based healthcare networks," *Neural Comput. Appl.*, 2024, doi: 10.1007/s00521-021-06441-5.
[17] J. A. Sultan Almotairi, Deepak Dasaratha Rao, Olayan Alharbi, Zaid Alzaid, Yasser M Hausawi, "Efficient Intrusion Detection using OptCNN-LSTM Model based on hybrid Correlation-based Feature Selection in IoMT.," *Fusion Pract. Appl.*, vol. 16, no. 1, 2024.
[18] M. Repetto, "Adaptive monitoring, detection, and response for agile digital service chains," *Comput. Secur.*, 2023, doi: 10.1016/j.cose.2023.103343.
[19] Olakunle Abayomi Ajala, Chuka Anthony Arinze, Onyeka Chrisanctus Ofodile, Chinwe Chinazo Okoye, and Obinna Donald Daraojimba, "Reviewing advancements in privacy-enhancing technologies for big data analytics in an era of increased surveillance," *World J. Adv. Eng. Technol. Sci.*, 2024, doi: 10.30574/wjaets.2024.11.1.0060.
[20] P. Bavadiya, P. Upadhyaya, A. C. Bhosle, S. Gupta, and N. Gupta, "AI-driven Data Analytics for Cyber Threat Intelligence and Anomaly Detection," in *2025 3rd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, 2025, pp. 677–681. doi: 10.1109/InCACCT65424.2025.11011329.
[21] M. Khule, D. Motwani, and D. Chauhan, "Adaptive Threat Intelligence: An Incremental Learning Approach for Detecting Evolving APT Attacks," in *2025 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, 2025, pp. 1–6. doi: 10.1109/AMATHE65477.2025.11081277.
[22] T. Swetha and S. Merakapudi, "Leveraging Transfer Learning for Enhanced Cybersecurity Threat Detection: A Novel Approach For Identifying Anomalies and Attacks," in *2025 International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, 2025, pp. 1–7. doi: 10.1109/ICKECS65700.2025.11035332.
[23] M. B. Vemula, K. P. Kumar, H. R. S. Thipparthi, and P. S. Jasti, "Network Anomaly Detection Using Generative Adversarial Networks," in *2024 First International Conference on Software, Systems and Information Technology (SSITCON)*, 2024, pp. 1–6. doi: 10.1109/SSITCON62437.2024.10796515.
[24] A. Kalra, K. S. Gill, M. Kumar, and R. Rawat, "CNN-Enhanced Network Security: A Sustainable Approach to Anomaly Detection," in *2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*, 2024, pp. 828–831. doi: 10.1109/ICAICCIT64383.2024.10912303.
[25] I. A. Kandhro *et al.*, "Detection of Real-Time Malicious Intrusions and Attacks in IoT Empowered Cybersecurity Infrastructures," *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3238664.
[26] F. Li and M. Liu, "A hybrid Convolutional and Recurrent Neural Network for Hippocampus Analysis in Alzheimer's Disease," *J. Neurosci. Methods*, 2019, doi: 10.1016/j.jneumeth.2019.05.006.
[27] R. Tarafdar and Y. Han, "Finding Majority for Integer Elements," *J. Comput. Sci. Coll.*, vol. 33, no. 5, pp. 187–191, 2018.
[28] S. Nokhwal, P. Chilakalapudi, P. Donekal, S. Nokhwal, S. Pahune, and A. Chaudhary, "Accelerating Neural Network Training: A Brief Review," in *2024 8th International Conference on Intelligent Systems Metaheuristics & Swarm Intelligence (ISMSI)*, New York, NY, USA: ACM, Apr. 2024, pp. 31–35. doi: 10.1145/3665065.3665071.
[29] F. S. Alrayes, M. Zakariah, S. U. Amin, Z. I. Khan, and J. S. Alqurni, "CNN Channel Attention Intrusion Detection System Using NSL-KDD Dataset," *Comput. Mater. Contin.*, vol. 79, no. 3, pp. 4319–4347, 2024, doi: 10.32604/cmc.2024.050586.
[30] A. D. Vibhute, C. H. Patil, A. V. Mane, and K. V. Kale, "Towards Detection of Network Anomalies using Machine Learning Algorithms on the NSL-KDD Benchmark Datasets," *Procedia Comput. Sci.*, vol. 233, no. January, pp. 960–969, 2024, doi: 10.1016/j.procs.2024.03.285.
[31] A. Tolba, N. N. Mostafa, and K. M. Sallam, "Hybrid Deep Learning-Based Model for Intrusion Detection," *Artif. Intell. Cybersecurity*, vol. 1, pp. 1–11, 2024, doi: 10.61356/j.aics.2024.1198.

